Notas de Probabilidade e Estatística

CASSIO ROBERTO DE MELO GODOI
Prof. de Matemática e Estatística
ESALQ/USP DE/IM/UFBA ESEB

USP – UFBA – ESEB
2001

Notas de Probabilidade e Estatística

1. Do que trata a Estatística.

O objeto de estudo da Estatística é ajudar ao pesquisador, com métodos descritivos como tabelas, gráficos e cálculos matemáticos, na busca de modelos novos ou explicações de fatos ocorridos no seu trabalho. Dessa forma podemos dizer que a Estatística está intimamente ligada à pesquisa científica ou qualquer outra forma de pesquisas, tais como pesquisa de mercado, pesquisa eleitoral e outras pesquisas. Convenhamos que não será fácil para alguém com nenhuma experiência com pesquisa compreender e saborear os métodos estatísticos.

O engenheiro-agrônomo estudando o melhoramento de milho necessita testar vários materiais genéticos e criar formas de cruzamentos de forma a criar variedades ou híbridos produtivos e resistentes às condições da região. Dessa forma esse especialista vai precisar da Estatística na análise dos experimentos de campo, estufa ou laboratório que utiliza. A Estatística apropriada para essa área está incluída na Genética Quantitativa. Os métodos estatísticos apropriados para a pesquisa nessa área pode mesmo se especializar em métodos para genética vegetal e para a genética animal.

O economista enfrenta problemas de ordem diferente e precisa normalmente coletar dados para sintetizar e mostrar uma situação social ou econômica. Pode querer criar modelos matemáticos para administradores públicos ou privados auxiliares na tomada de decisões. Lidam com problemas de engenharia econômica buscando formas ótimas de administração e controle. Os métodos estatísticos dessa área estão incluídos na Econometria e Pesquisa Operacional.

O médico trata de pesquisas diretamente aplicadas ao ser humano e a forma de planejamento, métodos estatísticos e medida dos riscos envolvidos é muito particular e incluem-se na área conhecida como Biometria.

Apesar da Estatística lidar com muitos cálculos matemáticos ela não é uma ciência baseada em axiomas, ou seja, dedutiva. Assim como a Física ela é essencialmente experimental e utiliza de certos princípios ou postulados que guiarão a obtenção de fórmulas úteis. Nesses princípios a Probabilidade será fundamental pois constitui-se no conceito apropriado para medir riscos. Portanto a Probabilidade é necessária para controlar os erros envolvidos nos métodos estatísticos aplicados a dados experimentais ou de levantamentos.

2. Dados estatísticos

Assim como na teoria dos conjuntos, uma **população** fica compreendida pelo conhecimento de seus elementos, aqui denominados genericamente de **indivíduos**. Para entender a amplitude desses conceitos tomemos alguns exemplos:

a) A ficha médica de uma pessoa contém informações de interesse do médico e do paciente. O conjunto de fichas constituem a população dos clientes do médico. Cada ficha corresponde efetivamente a um indivíduo. A informação contida nas fichas médicas podem ser de natureza estatística ou não. No caso da informação de particularidades pessoais e familiares o médico normalmente não usa de dados numéricos mas dados como altura, pressão sangüínea, peso, tipo de sangue e resultados de exames clínicos são de natureza estatística. Portanto o médico normalmente é, pela própria natureza de seu trabalho, um pesquisador e coleta diariamente dados estatísticos. Se ele vai usar os métodos estatísticos no seu trabalho é uma outra estória.

b) Um geneticista planta canteiros de milho e anota numa caderneta de campo, após certo tempo, o peso das espigas, número de espigas, altura da planta, composição química e física das sementes e outras variáveis de interesse do trabalho de pesquisa em andamento. Nesse caso o canteiro corresponde a indivíduo, aqui melhor denominado de unidade experimental e os possíveis canteiros de milho a população de interesse. Evidentemente quanto maior o escopo da pesquisa maior deve ser a diversidade de condições dos canteiros experimentais.

c) Jogam-se dois dados distintos, um vermelho e outro amarelo e observam-se e anotam-se certos resultados dessa jogada: resultado do dado vermelho, a soma dos pontos, a diferença entre o maior e o menor valor ocorrido, o máximo dos pontos obtidos, o mínimo dos pontos, a distância entre os dados em centímetros e outras possíveis variáveis de interesse. A população no caso serão todos as possíveis jogadas de dois dados e os indivíduos correspondem a cada um dos resultados com suas variáveis. Fazendo uma comparação com a ficha médica diz-se que a ficha de cada jogada de dois dados com as anotações dos resultados obtidos referentes às variáveis citadas constituem-se a unidade experimental.

d) Um economista registra semanalmente o comercio exterior de um segmento da economia anotando o volume de importações, o volume de exportações, o custo de transporte, a taxa de juro praticada nos negócios e demais variáveis de seu interesse. Essas fichas semanais configuram-se nas unidades experimentais e o conjunto de possíveis resultados semanais a população objeto do estudo.

O estudante já pode perceber a amplitude de situações em que surgem dados possíveis de serem tratados estatisticamente. Em resumo pode-se dizer que cada situação particular de pesquisa científica gera automaticamente dados estatísticos.

3. Tipos de dados estatísticos

Tomemos como exemplo uma caderneta de campo de um engenheiro-florestal interessado na pesquisa de qualidade de madeira para celulose e papel. Para seu trabalho esse engenheiro define como unidade experimental um talhão de um hectare. De cada talhão ele coleta dados referentes a

a) Altura média das árvores
b) Diâmetro à altura do peito (DAP)
c) Volume de madeira
d) Teor de lignina
e) Número de galhos pequenos
f) Susceptibilidade a nematóide
g) Qualidade da celulose
h) Cor da fibra
i) Resistência da fibra

O exemplo presente serve para ilustrar os vários tipos de dados usualmente encontrados. Esses dados podem classificar-se em **dados qualitativos** e **dados quantitativos**.

As variáveis f, g, e h são qualitativas e as demais quantitativas. A expressão de um dado qualitativo faz-se pela classificação do indivíduo numa de várias classes de uma categoria. São chamados também de dados categorizados. Temos no exemplo a categoria cor dividida em classes de cor. A susceptibilidade ao nematóide é outra categoria e positiva e negativa são suas classes. A qualidade da celulose é outra categoria e baixa, média e alta podem ser as classes.

As variáveis quantitativas podem ser discretas ou contínuas. As discretas são aquelas representadas por números inteiros como as contagens, pontuações etc. No exemplo o número de galhos é um dado inteiro. As contínuas são as representadas pelos números reais não enumeráveis, como peso, altura, porcentagem etc. No exemplo as variáveis a, b, c, e d são contínuas.

A importância dessa classificação de dados estatísticos deve-se ao fato de que o tratamento matemático dessas três grandes classes de dados diferem entre si a ponto de serem tratados separadamente na literatura especializada.

4. Estratégia de ensino da Estatística

Vamos obedecer um dos mandamentos principais de Descartes em seu Discurso do Método ou seja que deve-se sempre ao abordar um problema começar pelo concreto para depois passar ao abstrato. Do particular para o geral. Essa estratégia se bem usada normalmente é mais didática e mais objetiva e é a que adotaremos neste curso. Por exemplo, em seguida vamos abordar a análise de dados categorizados de onde nasce naturalmente a idéia de probabilidade e o conceito de independência estocástica ou probabilística.

5. Dados Categorizados

População, Amostra, Eventos, Freqüência relativa e Probabilidade.

Seguindo a orientação cartesiana comecemos com um exemplo (**Mosteller, Rourke e Thomas**): uma prefeitura elaborou uma pesquisa cujo questionário de abordagem continha, dentre outras, as seguintes perguntas:

1. A respeito da lei em discussão na câmara de vereadores proibindo barcos a motor na represa da cidade o(a) senhor(a):
   a) Concorda com a lei.
   b) Discorda da lei.
   c) Sem comentários.

2. Quanto a posse de barcos o(a) senhor(a):

   a) Possui só barco a motor.
   b) Possui só barco a vela.
   c) Possui barco a motor e barco a vela.
   d) Não possui nem barco a vela nem a motor.

Nesse exemplo temos duas categoria de dados: a categoria "***sobre a lei***" com três classes de respostas e a categoria "***posse de barco***" com quatro classes de respostas. São dados obtidos de múltiplas escolhas, como é usualmente conhecidos.
A forma de múltipla escolha é a mais usual nas pesquisas de opinião de modo geral e o conjunto das perguntas visa a coleta de informações necessárias para o diagnóstico de uma situação em estudo.

Se a pesquisa for feita sobre todos indivíduos envolvidos no problema temos um **censo**, caso contrário se os questionários limitarem-se a uma parte dos indivíduos temos uma **amostragem**. Ao conjunto de todos os indivíduos chamaremos de **População** e qualquer parte dessa população chamaremos de **Amostra**.

6. Tabela de Contingência ou de Dependência das Categorias

Uma tabela estatística é um resumo dos resultados de uma pesquisa disposta na forma retangular de linhas e colunas. Dispomos nas linhas as classes de uma das categorias e nas colunas as classes da outra categoria. No exemplo citado temos o seguinte resultado:

Opinião de moradores sobre a lei de proibição de barco a motor em 1960

Classe	Motor	Vela	Vela e Motor	Nenhum	Total
A favor da lei	0	7	1	18	26
Contra a lei	20	2	3	5	30
Sem opinião	0	1	1	12	14
Total	20	10	5	35	70

Fonte: Mosteller, Rourke e Thomas: Probability with Statistical Applications

Uma tabela estatística divide-se em três partes: cabeçalho, corpo e rodapé. O cabeçalho responde as questões: **Quê?, Onde?** e **Quando?** O corpo expressa em números o cruzamento de duas classes das respectivas categorias em estudo e o rodapé completa informações adicionais diversas, como a fonte dos dados.

Os números da tabela desse exemplo são chamados **freqüências brutas**, os totais são conhecidos como **freqüências marginais** da categoria e o total é conhecido por **tamanho da amostra** ou **tamanho da população** conforme for o caso.

Dividindo-se cada célula da tabela pelo total 70 obtemos as freqüências relativas correspondentes:

Opinião de moradores sobre a lei de proibição de barco a motor em 1960

Classe	Motor	Vela	Vela e Motor	Nenhum	Total
A favor da lei	0	7/70	1/70	18/70	26/70
Contra a lei	20/70	2/70	3/70	5/70	30/70
Sem opinião	0	1/70	1/70	12/70	14/70
Total	20/70	10/70	5/70	35/70	1

Fonte: Mosteller, Rourke e Thomas: Probability with Statistical Applications

Caso a tabela fosse da população e não apenas da amostra as freqüências relativas passariam a chamar-se **probabilidades**.

7. Probabilidade

Probabilidade é uma função matemática definida para **eventos ou fatos** de um **experimento**. Quando um pesquisador realiza uma pesquisa ele estará fazendo um experimento ou experiência. Os fatos de uma experiência não podem ser antecipados com certeza e dizemos que os eventos decorrentes de uma experiência são **aleatórios**. O limite da freqüência relativa de um dado evento quando a amostra tende para a população, se existir, é chamado de **probabilidade** desse evento.

Como na teoria dos conjuntos há uma álgebra de eventos:

a) Se pelo menos um dos eventos **A** ou **B** ocorrem dizemos que **A**$\cup$**B** ocorre.
b) Se **A** não ocorre então dizemos que **~A** ocorre.
c) Se ambos **A** e **B** ocorrem então dizemos que **A**$\cap$**B** ocorre.
d) Se um evento ocorre com certeza então chamamos de **Ω** ou evento universo**.**
e) Se um evento não pode ocorrer chamamos de ∅ ou evento vazio.

Dizemos que A e B são eventos **contraditórios** ou **incompatíveis** ou **mutuamente exclusivos** ou **disjuntos** se a ocorrência de A se realiza com a não ocorrência de B e vice-versa, a ocorrência de B determina necessariamente a não ocorrência de A. .Às vezes para enfatizar que os eventos são incompatíveis alguns autores escrevem A+B no lugar na terminologia mais ampla A$\cup$B. Na tabela acima, como em qualquer tabela de eventos, há uma partição do universo em 12 eventos incompatíveis a saber:

$L \cap M$	$L\cap V$	$L\cap MV$	$L\cap N$
$C\cap M$	$C\cap V$	$C\cap MV$	$C\cap N$
$S\cap M$	$S\cap V$	$S\cap MV$	$S\cap N$

Onde L, C e S representam “Pela lei”, “Contra a lei” e “Sem comentário” e M, V, MV e N representam os eventos “a motor”, “à vela”, “a motor e à vela” e “nenhum”. A ocorrência de qualquer um desses 12 eventos pressupõe a não ocorrência dos demais. Podemos então escrever:

$$\begin{aligned}\Omega = \; & L \cap M + L\cap V + L\cap MV + L\cap N + \\ & C\cap M + C\cap V + C\cap MV + C\cap N + \\ & S\cap M + S\cap V + S\cap MV + S\cap N\end{aligned}$$

$$L = L \cap M + L\cap V + L\cap MV + L\cap N$$
$$C = C\cap M + C\cap V + C\cap MV + C\cap N$$
$$S = S\cap M + S\cap V + S\cap MV + S\cap N$$

$$\Omega = L + C + S$$

M	=	$M\cap L$	+	$M\cap C$	+	$M\cap S$
V	=	$V\cap L$	+	$V\cap C$	+	$V\cap S$
MV	=	$MV\cap L$	+	$MV\cap C$	+	$MV\cap S$
N	=	$N\cap L$	+	$N\cap C$	+	$N\cap S$

$$\Omega = M + V + MV + N$$

Reservamos o nome **partição** para a união disjunta de eventos. Da primeira das equações acima decorre que qualquer possível indivíduo da pesquisa cairá numa e apenas numa das 12 células da tabela de contingência. Na amostra observada 0 indivíduos só com barco a motor disseram sim à lei, 7/70 = 1/10 dos indivíduos só com barco a vela disseram sim à lei e assim por diante.

Os eventos L, C e S são eventos que se limitam a aplicação da lei, ignorando o tipo de barco que o indivíduo tenha e quando queremos enfatizar esse caráter dizemos que são eventos marginais. Nesse sentido os eventos M, V, MV e N são também eventos marginais. Em contra partida os eventos como $L\cap M$ etc. são chamados eventos conjuntos pois são eventos produzidos pela ocorrência simultânea de dois eventos.

8. A Função de Probabilidade

Seja E um experimento e A e B dois eventos quaisquer definidos nesse experimento. A função matemática P tal que:

1) $P(A) \geq 0$, para todo A
2) $P(\Omega) = 1$
3) $P(A_1 \cup A_2 \cup A_3 \ldots) = P(A_1) + P(A_2) + P(A_3) + \ldots$, sempre que $A_1, A_2, A_3, \ldots$ forem eventos incompatíveis

Podemos perceber no exemplo acima que a freqüência relativa segue perfeitamente as regras da função de probabilidade. Na verdade se a população de moradores se restringisse aos 70 indivíduos pesquisados a freqüência relativa de eventos ou combinações de eventos torna-se a função de probabilidade. Caso estejamos lidando com uma amostra então dizemos que a freqüência relativa estima a verdadeira probabilidade. No exemplo escrevemos f_{ij} simbolizando a freqüência referente à linha **i** e à coluna **j** e escrevemos p_{ij} para a probabilidade respectiva. Se **n** representar o total de indivíduos pesquisados então n p_{ij} será a freqüência bruta da célula da linha **i**, coluna **j**.

Exercício:

a) Qual a freqüência relativa do evento W= {“ Morador com barco a motor”}?

b) Qual a freqüência relativa do evento V= {" Morador com barco a vela"}?

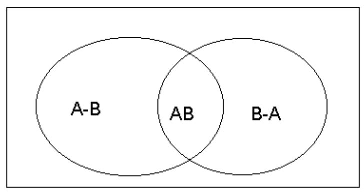
c) Qual a freqüência relativa do evento $W \cap V$?

d) Qual a freqüência relativa do evento $W \cup V$?

c) Mostre que no exemplo $f\{W \cup V) = f\{W\} + f\{V\} - f\{W \cap V\}$.

A diferença **A-B** é definida pela expressão **A∩~B** (o indivíduo pertence à A mas não à B). Dessa forma podemos partir o evento união $A \cup B$ em $(A-B)+(A \cap B)+(B-A)$. O diagrama de Venn ilustra bem esse fato:

Diagrama de Venn de A e B



onde AB significa $A \cap B$. Como os eventos são incompatíveis podemos escrever, devido aos axiomas da Probabilidade, que

$$
\begin{aligned}
P(A \cup B) &= P(A-B)+P\{A \cap B)+P(B-A) \\
P(A) &= P(A-B)+P(A \cap B) \\
P(B) &= P(B-A)+P(A \cap B) \\
P(A)+P(B) &= P(A-B)+P(A \cap B)+ P(B-A)+P(A \cap B) \\
P(A)+P(B) &= P(A \cup B)+P(A \cap B)
\end{aligned}
$$

Mostrando que em geral $P(A \cup B)=P(A)+P(B)-P(A \cap B)$. Somente quando os eventos são incompatíveis temos que $P(A \cap B)=0$, donde $P(A \cup B)=P(A)+P(B)$.

Qualquer resultado obtido para probabilidades serve imediatamente para freqüências relativas.

Eventos do tipo "pelo menos um de vários eventos" corresponde sempre a uma união de eventos e é necessário cuidado no cálculo de sua probabilidade, pois é necessário cuidar dos eventos conjuntos ou eventos de interseção.

Se tivermos 3 eventos A, B e C então podemos escrever

$$
\begin{aligned}
P(A \cup B \cup C) &= P([A \cup B] \cup C)=P(A \cup B)+P(C)-P([A \cup B] \cap C) \\
&= P(A)+P(B)-P(A \cap B)+P(C)-P([A \cap C] \cup [B \cap C])
\end{aligned}
$$

Mas, $P([A\cap C]\cup[B\cap C]) = P(A\cap C)+ P(B\cap C)-P(A\cap B\cap C)$

Donde,

$$P(A\cup B\cup C) = P(A)+P(B)-P(A\cap B)+P(C)- [P(A\cap C)+ P(B\cap C)-P(A\cap B\cap C)]$$

Ou ainda,

$$P(A\cup B\cup C) = P(A)+P(B)+P(C)-P(A\cap B)-P(A\cap C)- P(B\cap C)+P(A\cap B\cap C)$$

No caso de três eventos para que possamos obter a probabilidade da união pela soma das probabilidades desses eventos é necessário que todos os eventos conjuntos dois a dois ou três a três tenham probabilidades nulas.

9. Freqüências e probabilidades condicionais

No exemplo dado se restringirmos a pesquisa apenas às pessoas que opinaram isso vai alterar as freqüências relativas dos eventos,

Classe	Motor	Vela	Vela e Motor	Nenhum	Total
A favor da lei	0	7/56	1/56	18/56	26/56
Contra a lei	20/56	2/56	3/56	5/56	30/56
Total	20/56	9/56	4/56	23/56	1

Temos uma nova função de probabilidade relacionada à primeira, pois foi obtida da tabela anterior. O conhecimento prévio de um certo evento pode portanto modificar nosso grau de informação a respeito de um evento.

Seja E um evento onde $P(E)\neq 0$ e A outro evento do experimento. Então a **probabilidade condicional dado E** de A será definida pela expressão:

$$P(A|E)=P(A\cap E)/P(E) \quad \text{ou} \quad P_E(A)= P(A\cap E)/P(E)$$

Notem que o símbolo | é diferente da barra de divisão / e realmente tem significado muito diferente. A linguagem matemática deve ser precisa e os símbolos devem ser inequívocos. O Símbolo |E lê-se **dado E** enquanto / lê-se **dividido por**.

Complementando diríamos que um experimento associa-se a muitas funções probabilísticas dependendo do grau de conhecimento que temos dos eventos que nele podem ocorrer. Uma informação adicional pode alterar muito as probabilidades dos eventos do experimento.

Transportado para freqüências relativas essa definição estende-se de forma natural e chamamos de freqüências relativas condicionais. Considerando apenas os 70 moradores como população alvo e fixando as classes da categoria **Lei** teremos três funções de probabilidades condicionais a saber:

Classe	Motor	Vela	Vela e Motor	Nenhum	Total
A favor da lei	0	7/26	1/26	18/26	1
Contra a lei	20/30	2/30	3/30	5/30	1
Sem opinião	0	1/14	1/14	12/14	1

obtida calculando-se a primeira linha assim:

$$P(M|L) = P(M\cap L)/P(L) = (0/70)/(26/70)$$

$$P(V|L) = P(V\cap L)/P(L) = (7/70)/(26/70)$$

$$P(MV|L) = P(MV\cap L)/P(L) = (1/70)/(26/70)$$

$$P(N|L) = P(N\cap L)/P(L) = (18/70)/(26/70)$$

Exercício: Verifiquem se as demais linhas estão corretas.

Se fixarmos as classes da categoria **Barco** teremos mais quatro funções de probabilidades condicionais a saber:

Classe	Motor	Vela	Vela e Motor	Nenhum
A favor da lei	0	7/10	1/5	18/35
Contra a lei	20/20	2/10	3/5	5/35
Sem opinião	0	1/10	1/5	12/35
Total	1	1	1	1

Obtida calculando-se a segunda coluna assim:

$$P(L|V) = P(L\cap V)/P(V) = (7/70)/(10/70)$$

$$P(C|V) = P(C\cap V)/P(V) = (2/70)/(10/70)$$

$$P(S|V) = P(S\cap V)/P(V) = (1/70)/(10/70)$$

Do primeiro conjunto de condicionais podemos perceber como o conhecimento prévio de uma determinada classe da categoria **Lei** muda as probabilidades dos eventos.

É importante entender como reconhecer uma função de probabilidade ou também chamada **função de distribuição de probabilidades** (f.d.p). Nesse exemplo, a partir de uma mesma pesquisa deduzimos várias outras distribuições de probabilidades. Se soubermos que um morador é a favor da lei a probabilidade de Ter apenas barco a motor é 0/26=0 mas se soubermos que um morador é contra a lei então a probabilidade de ter apenas barco a motor passa a ser de 20/30=2/3.

10. Um exemplo bem matemático: Jogada de dois dados: **Introdução à Variáveis Aleatórias**.

Considere o experimento da jogada de dois dados, um vermelho e outro preto. Se tivermos uma categoria de resultado com classes definidas por números reais então estamos diante de uma **variável aleatória**. Por exemplo neste experimento poderíamos estar interessados em:

a) X : No. de pontos do dado preto
b) Y : No. de pontos do dado vermelho.
c) S = X+Y
d) D = X-Y
e) U = Max (X,Y)
f) V = Min (X,Y)

Tabela 1. Resultados possíveis das variáveis aleatórias X, Y, S, D, U e V

x	y	s	d	u	v	Probabilidade
1	1	2	0	1	1	1/36
1	2	3	-1	2	1	1/36
1	3	4	-2	3	1	1/36
1	4	5	-3	4	1	1/36
1	5	6	-4	5	1	1/36
1	6	7	-5	6	1	1/36
2	1	3	1	2	1	1/36
2	2	4	0	2	2	1/36
2	3	5	-1	3	2	1/36
2	4	6	-2	4	2	1/36
2	5	7	-3	5	2	1/36
2	6	8	-4	6	2	1/36
3	1	4	2	3	1	1/36
3	2	5	1	3	2	1/36
3	3	6	0	3	3	1/36
3	4	7	-1	4	3	1/36
3	5	8	-2	5	3	1/36
3	6	9	-3	6	3	1/36
4	1	5	3	4	1	1/36
4	2	6	2	4	2	1/36
4	3	7	1	4	3	1/36
4	4	8	0	4	4	1/36
4	5	9	-1	5	4	1/36
4	6	10	-2	6	4	1/36
5	1	6	4	5	1	1/36
5	2	7	3	5	2	1/36
5	3	8	2	5	3	1/36
5	4	9	1	5	4	1/36
5	5	10	0	5	5	1/36
5	6	11	-1	6	5	1/36
6	1	7	5	6	1	1/36
6	2	8	4	6	2	1/36
6	3	9	3	6	3	1/36
6	4	10	2	6	4	1/36
6	5	11	1	6	5	1/36
6	6	12	0	6	6	1/36

Um jogo de dados “honestos” é um experimento tão simples que não será preciso nem dados estatísticos para conhecermos as probabilidades dos eventos associados a tal experimento. Podemos então dizer que é possível criar um **modelo estocástico** ou **modelo probabilístico** desse experimento. Para tal necessitamos apenas criar uma função de probabilidades aceitável. Consideremos os possíveis resultados (X,Y) desse experimento contidos no universo da pesquisa :

$$\Omega = \{ \begin{array}{l} (1,1), (1,2), (1,3), (1,4), (1,5), (1,6), \\ (2,1), (2,2), (2,3), (2,4), (2,5), (2,6), \\ (3,1), (3,2), (3,3), (3,4), (3,5), (3,6), \\ (4,1), (4,2), (4,3), (4,4), (4,5), (4,6), \\ (5,1), (5,2), (5,3), (5,4), (5,5), (5,6), \\ (6,1), (6,2), (6,3), (6,4), (6,5), (6,6) \} \end{array}$$

e vamos atribuir 1/36 como a probabilidade de qualquer resultado de Ω. Podemos demonstrar que essa definição satisfaz aos 3 axiomas citados e podemos usar esse modelo para calcular a probabilidade de qualquer evento do experimento da jogada de dois dados honestos.

Deduzir a **distribuição de probabilidades** de uma variável aleatória **X** é obter seus possíveis valores (**espaço amostral de X**) e suas respectivas probabilidades. Para obter as distribuições de X, Y, S, D, U e V podemos organizar a tabela auxiliar vista acima, onde as seis primeiras colunas referem-se às variáveis aleatórias de interesse e a sétima coluna refere-se à probabilidade dos resultados possíveis. Cada uma das 36 linhas refere-se a um resultado possível em particular Decorre dessa tabela auxiliar a distribuição de todas as seis variáveis. A saber:

x	P(x)	y	P(y)	u	P(u)	v	P(v)
1	6/36	1	6/36	1	1/36	1	11/36
2	6/36	2	6/36	2	3/36	2	9/36
3	6/36	3	6/36	3	5/36	3	7/36
4	6/36	4	6/36	4	7/36	4	5/36
5	6/36	5	6/36	5	9/36	5	3/36
6	6/36	6	6/36	6	11/36	6	1/36
Total	1		1		1		1

s	P(s)	d	P(d)
2	1/36	-5	1/36
3	2/36	-4	2/36
4	3/36	-3	3/36
5	4/36	-2	4/36
6	5/36	-1	5/36
7	6/36	0	6/36
8	5/36	1	5/36
9	4/36	2	4/36
10	3/36	3	3/36
11	2/36	4	2/36
12	1/36	5	1/36
Total			1

Exercício: Considerando-se o experimento dos dois dados, obtenha a distribuição das variáveis aleatórias:

a) $L = 2X + Y$ b) $Q = XY$

11. Teorema do Produto de Probabilidades

Da expressão $P(A|E)=P(A\cap E)/P(E)$ resulta que

$$P(A\cap E)= P(A|E)\, P(E),$$

uma fórmula para obter-se a probabilidade da ocorrência conjunta de dois eventos. Para três eventos teremos

$$P(A\cap E\cap M) = P(A\cap E|M)\, P(M).$$

Mas aplicando essa fórmula recursivamente teremos

$$P(A\cap E|M) = P(A | M\cap E)P(E|M) \text{ ou } P_M(A\cap E)=P_{M\cap E}(A)P_M(E)$$

Resultando

$$P(A\cap E\cap M) = P(A | E\cap M)\, P(E|M)\, P(M)$$

donde valerão também as demais cinco expressões semelhantes ao mudar a ordem dos eventos:

$$P(A\cap E\cap M) = P(A | E\cap M)\, P(M|E)\, P(E)$$

$$P(A\cap E\cap M) = P(E | A\cap M)\, P(A|M)\, P(M)$$

$$P(A\cap E\cap M) = P(E | A\cap M)\, P(M|A)\, P(A)$$

$$P(A\cap E\cap M) = P(M | A\cap E)\, P(A|E)\, P(E)$$

$$P(A\cap E\cap M) = P(M | A\cap E)\, P(E|A)\, P(A)$$

Exercício: O aluno deve tentar escrever uma das fórmulas da probabilidade da ocorrência simultânea de quatro eventos A, B, C e D. Quantas fórmulas equivalentes teríamos?

a) Considere uma caixa com 9 bolas sendo 4 brancas, 3 pretas e 2 vermelhas. Um experimento consiste em retirar em seqüência 3 bolas, sem reposição dessas bolas à caixa. Qual a probabilidade de sair uma seqüência de bolas de cores distintas?

Nomeemos os eventos "saída de branca", "saída de preta" e "saída de vermelha" por B, N e V e indiquemos a numeração das letras como sendo a ordem de retirada da bola. Chamemos de T o evento saída de bolas com três cores distintas. Então as seis seqüências componentes de T podem ser expressas pela equação

$$T = (B_1 \cap N_2 2 \cap V_3) + (B_1 \cap V_2 \cap N_3) + (N_1 \cap B_2 \cap V_3) + (N_1 \cap V_2 \cap B_3) + (V_1 \cap B_2 \cap N_3) + (V_1 \cap N_2 \cap B_3)$$

onde, por exemplo, $B_1 \cap V_2 \cap N_3$ significa "sair branca na primeira retirada" **e** "sair vermelha na segunda retirada" **e** "sair preta na terceira retirada" e a soma de eventos indicando a união de eventos incompatíveis, pois qualquer das seqüências só pode ocorrer se as demais não ocorrerem. Pelo axioma 3 podemos escrever:

$$P(T)=P(B_1\cap N_2 2\cap V_3)+P(B_1\cap V_2\cap N_3)+P(N_1\cap B_2\cap V_3)+P(N_1\cap V_2\cap B_3)+P(V_1\cap B_2\cap N_3)+P(V_1\cap N_2\cap B_3)$$

Calculando os componentes da soma pela probabilidade da interseção de três conjuntos, temos:

$$P(B_1\cap N_2\cap V_3)=P(B_1)P(N_2\mid B_1)P(V_3\mid B_1\cap N_2)$$

$$=\frac{4}{9}\times\frac{3}{8}\times\frac{2}{7}=\frac{1}{21}$$

$$P(B_1\cap V_2\cap N_3)=P(B_1)P(V_2\mid B_1)P(N_3\mid B_1\cap V_2)$$

$$=\frac{4}{9}\times\frac{2}{8}\times\frac{3}{7}=\frac{1}{21}$$

$$P(N_1\cap B_2\cap V_3)=P(N_1)P(B_2\mid N_1)P(V_3\mid N_1\cap B_2)$$

$$=\frac{3}{9}\times\frac{4}{8}\times\frac{2}{7}=\frac{1}{21}$$

como os demais componentes correspondem a permutações dos numeradores, resulta que

$$P(T)=6\times\frac{1}{21}=\frac{6}{21}=\frac{2}{7}$$

Exercício: O estudante nesse ponto deve tentar calcular a probabilidade de nesse experimento sair

a) A = “branca como segunda bola”
b) B = “exatamente duas bolas pretas”
c) $A\cap B$
d) $A\cup B$

12. Distribuição Conjunta de Variáveis Aleatórias: Distribuições Conjuntas, Distribuições Marginais e Distribuições Condicionais.

Considere o experimento da jogada de dois dados. Vimos como obter as distribuições das variáveis individualmente. Em seqüência vamos obter a **distribuição conjunta** de pares de variáveis aleatórias.

Tomemos, por exemplo, o par S e D. A distribuição conjunta de **S e D** pode ser obtida da tabela 1 somando-se as probabilidades dos possíveis pares de valores de S e D, obtendo-se a tabela 2 seguinte:

Tabela 2. Distribuição Conjunta de S e D

d / s	-5	-4	-3	-2	-1	0	1	2	3	4	5	P(s)
2	0	0	0	0	0	1/36	0	0	0	0	0	1/36
3	0	0	0	0	1/36	0	1/36	0	0	0	0	2/36
4	0	0	0	1/36	0	1/36	0	1/36	0	0	0	3/36
5	0	0	1/36	0	1/36	0	1/36	0	1/36	0	0	4/36
6	0	1/36	0	1/36	0	1/36	0	1/36	0	1/36	0	5/36
7	1/36	0	1/36	0	1/36	0	1/36	0	1/36	0	1/36	6/36
8	0	1/36	0	1/36	0	1/36	0	1/36	0	1/36	0	5/36
9	0	0	1/36	0	1/36	0	1/36	0	1/36	0	0	4/36
10	0	0	0	1/36	0	1/36	0	1/36	0	0	0	3/36
11	0	0	0	0	1/36	0	1/36	0	0	0	0	2/36
12	0	0	0	0	0	1/36	0	0	0	0	0	1/36
P(d)	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36	1

Podemos obter as **distribuições condicionais** de **S dado D** e **D dado S** (tabelas 3 e 4).

Tabela 3. Distribuição Condicional de S dado D

d / s	-5	-4	-3	-2	-1	0	1	2	3	4	5	P(s)
2	0	0	0	0	0	1/6	0	0	0	0	0	1/36
3	0	0	0	0	1/5	0	1/5	0	0	0	0	2/36
4	0	0	0	1/4	0	1/6	0	1/4	0	0	0	3/36
5	0	0	1/3	0	1/5	0	1/5	0	1/3	0	0	4/36
6	0	1/2	0	1/4	0	1/6	0	1/4	0	1/2	0	5/36
7	1	0	1/3	0	1/5	0	1/5	0	1/3	0	1	6/36
8	0	1/2	0	1/4	0	1/6	0	1/4	0	1/2	0	5/36
9	0	0	1/3	0	1/5	0	1/5	0	1/3	0	0	4/36
10	0	0	0	1/4	0	1/6	0	1/4	0	0	0	3/36
11	0	0	0	0	1/5	0	1/5	0	0	0	0	2/36
12	0	0	0	0	0	1/6	0	0	0	0	0	1/36
Total	1	1	1	1	1	1	1	1	1	1	1	1

Tabela 4. Distribuição Condicional de D dado S											
d s	-5	-4	-3	-2	-1	0	1	2	3	4	5	Total
2	0	0	0	0	0	1	0	0	0	0	0	1
3	0	0	0	0	1/2	0	1/2	0	0	0	0	1
4	0	0	0	1/3	0	1/3	0	1/3	0	0	0	1
5	0	0	1/4	0	1/4	0	1/4	0	1/4	0	0	1
6	0	1/5	0	1/5	0	1/5	0	1/5	0	1/5	0	1
7	1/6	0	1/6	0	1/6	0	1/6	0	1/6	0	1/6	1
8	0	1/5	0	1/5	0	1/5	0	1/5	0	1/5	0	1
9	0	0	1/4	0	1/4	0	1/4	0	1/4	0	0	1
10	0	0	0	1/3	0	1/3	0	1/3	0	0	0	1
11	0	0	0	0	1/2	0	1/2	0	0	0	0	1
12	0	0	0	0	0	1	0	0	0	0	0	1
P(d)	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36	1

As tabelas 3 e 4 permitem responder perguntas como P(D=3|S=4), P(S>4|D=2) etc. Observando os dois conjuntos de distribuições condicionais deduzimos que a distribuição da soma S pode se alterar muito dependendo do conhecimento ou não do valor da diferença D. Essa falta de uniformidade das distribuições condicionais é que determina o conceito de **dependência probabilística entre as variáveis aleatórias S e D**.

Tabela 5 Distribuição conjunta de X e Y							
y x	1	2	3	4	5	6	P(x)
1	1/36	1/36	1/36	1/36	1/36	1/36	1/6
2	1/36	1/36	1/36	1/36	1/36	1/36	1/6
3	1/36	1/36	1/36	1/36	1/36	1/36	1/6
4	1/36	1/36	1/36	1/36	1/36	1/36	1/6
5	1/36	1/36	1/36	1/36	1/36	1/36	1/6
6	1/36	1/36	1/36	1/36	1/36	1/36	1/6
P(y)	1/6	1/6	1/6	1/6	1/6	1/6	1

Tabela 6 Distribuições condicionais de X dados Y=1, Y=2, ... , Y=6							
y x	1	2	3	4	5	6	P(x)
1	1/6	1/6	1/6	1/6	1/6	1/6	1/6
2	1/6	1/6	1/6	1/6	1/6	1/6	1/6
3	1/6	1/6	1/6	1/6	1/6	1/6	1/6
4	1/6	1/6	1/6	1/6	1/6	1/6	1/6
5	1/6	1/6	1/6	1/6	1/6	1/6	1/6
6	1/6	1/6	1/6	1/6	1/6	1/6	1/6
Total	1	1	1	1	1	1	1

Tabela 7 Distribuições condicionais de Y dados X=1, X=2, ... , X=6							
y / x	1	2	3	4	5	6	Total
1	1/6	1/6	1/6	1/6	1/6	1/6	1
2	1/6	1/6	1/6	1/6	1/6	1/6	1
3	1/6	1/6	1/6	1/6	1/6	1/6	1
4	1/6	1/6	1/6	1/6	1/6	1/6	1
5	1/6	1/6	1/6	1/6	1/6	1/6	1
6	1/6	1/6	1/6	1/6	1/6	1/6	1
P(y)	1/6	1/6	1/6	1/6	1/6	1/6	1

Em contrapartida as tabelas 5 e 6 indicam que as distribuições condicionais de X dado Y=y são idênticas à marginal (ou incondicional) de X e as distribuições condicionais de Y dado X=x são idênticas à marginal (ou incondicional) de Y, definindo X e Y como **variáveis aleatórias independentes** e nesse caso temos que $P(X = x, Y = y) = P(X = x) \times P(Y = y)$ para todo par (x,y).

13. Teorema da Probabilidade Total e Teorema de Bayes

Consideremos $\Omega = P_1 + P_2 + \cdots + P_t$ uma partição do espaço amostral, significando como já vimos uma união exaustiva e mutuamente exclusiva. Considere E um evento qualquer de Ω. Então

$$P(B) = P\{E \cap (P_1 + P_2 + \cdots + P_t)\} = P(E \cap P_1) + P(E \cap P_2) + \cdots + P(E \cap P_t)$$

$$P(E) = P(P_1) \cdot P(E|P_1) + P(P_2) \cdot P(E|P_2) + \cdots + P(P_t) \cdot P(E|P_t)$$

portanto a probabilidade de um evento E pode ser obtida pela média ponderada das probabilidades condicionais desse evento E dadas cada uma das condições definidas pela partição realizada. Esse resultado é chamado de **teorema da probabilidade total**. Tomemos um exemplo ilustrativo: temos 2 duas urnas: U_1 com 6 bolas brancas e 4 vermelhas e U_2 com 4 bolas brancas e 6 vermelhas. Sorteamos uma urna e tiramos uma bola. Qual a probabilidade de sair uma bola branca? Seja E o evento "sair bola branca", P_1 o evento "escolher urna 1"e P_2 o evento "escolher urna 2". Pela mesma construção acima descrita teremos $\Omega = P_1 + P_2$, $P(P_1) = P(P_2) = 1/2$, $P(E|P_1) = 6/10$ e $P(E|P_2) = 4/10$, donde

$$P(E) = \frac{6}{10} \cdot \frac{1}{2} + \frac{4}{10} \cdot \frac{1}{2} = \frac{1}{2}$$

Modifiquemos o exemplo de forma a escolher a urna 1 caso a jogada de um dado regular saia 5 ou 6 e escolher a urna 2 nos outros casos. Nesse caso $P(P_1) = 1/3$ e $P(P_2) = 2/3$ resultando que

$$P(E) = \frac{6}{10} \cdot \frac{1}{3} + \frac{4}{10} \cdot \frac{2}{3} = \frac{6+8}{30} = \frac{14}{30} = \frac{7}{15}$$

A fórmula da probabilidade total permite o cálculo de probabilidades de eventos em etapas o que em regra geral facilita o raciocínio.

Outros exemplos de aplicação do teorema da probabilidade total constam das tabelas de distribuições condicionais de S dado D e D dado S. Podemos observar que P(s) é igual à média das probabilidades condicionais P(S=s|D=d) ponderadas pelas probabilidades marginais P(d). De forma análoga P(d) é igual à média das probabilidades condicionais P(D=d|S=s) ponderadas pelas probabilidades marginais P(s).

Consideremos o problema geral em que os dados conhecidos sejam

$$P(E|P_1),\ P(E|P_2),\ ...,\ P(E|P_t)$$
$$P(P_1),\ P(P_2),\ ...\ P(P_t),$$

e desejamos calcular $P(P_1|E)$, $P(P_2|E)$, ..., $P(P_t|E)$. Poderíamos interpretar esse problema clássico como aquele em que são dados as probabilidades de um certo evento E sob várias condições P_1, P_2, ... , P_t e quer-se obter as probabilidades das condições dada a ocorrência do evento E. Esse é chamado de problema bayesiano ou problema da probabilidade das hipóteses ou ainda problema da inversão das probabilidades condicionais.

Usando o teorema da probabilidade total teremos

$$P(P_i \mid E) = \frac{P(P_i \cap E)}{P(E)}$$

$$P(P_i \mid E) = \frac{P(P_i) \cdot P(E \mid P_i)}{P(P_1) \cdot P(E|P_1) + P(P_2) \cdot P(E|P_2) + \cdots + P(P_t) \cdot P(E|P_t)}$$

para $i = 1, 2, \cdots, t$

A fórmula acima é conhecida como fórmula de Bayes.

Exemplo: Uma doença atinge 5% de uma população. Um pesquisador inventou um teste clínico com as seguintes características: o teste dá negativo em 80% dos indivíduos isentos da doença e dá positivo para 90% dos indivíduos efetivamente doentes. Portanto as probabilidades dos dois tipos de erros são, respectivamente, 20% e 10%. Chamando de S o evento “Indivíduo é sadio” , de D o evento “Indivíduo é doente”, + o evento “Teste positivo” e – o evento “Teste negativo” teremos, na linguagem probabilística:

$$P(-\mid S)=0.80 \qquad P(+\mid S)=0.20 \qquad P(+\mid D)=0.90 \qquad P(-\mid D)=0.10$$

Um indivíduo dessa população tomado ao acaso realiza o teste e, é claro, não se sabe se esse indivíduo é doente ou sadio. A pergunta é: se o teste realizado dá positivo, qual a probabilidade desse indivíduo ser realmente doente?. A pergunta complementar é : se o teste foi negativo, qual a probabilidade desse indivíduo ser sadio? Matematicamente dizendo quer-se $P(D\mid +)$ e $P(S\mid -)$. Esse como inúmeros outros ilustram a aplicação da fórmula de Bayes.

Solução: Ω=D+S, Na primeira questão o evento + toma o lugar de E e na Segunda questão – toma o lugar de E na fórmula geral. Assim temos:

$$P(D\mid +)=\frac{P(D)\cdot P(+\mid D)}{P(D)\cdot P(+\mid D)+P(S)\cdot P(+\mid S)} = \frac{0.05\cdot 0.90}{0.05\cdot 0.90+0.95\cdot 0.20} = 0{,}1915$$

$$P(S\mid -)=\frac{P(S)\cdot P(-\mid S)}{P(D)\cdot P(-\mid D)+P(S)\cdot P(-\mid S)} = \frac{0.95\cdot 0.80}{0.05\cdot 0.10+0.95\cdot 0.80} = 0{,}9935$$

Esse resultado deve surpreender a muitos pois um teste aparentemente “bom” mostra-se um teste frágil ao indicar somente 19,15% dos indivíduos como doentes quando acusa resultado positivo ou, de uma outra forma, 81,85% dos indivíduos positivos para o teste serão sadios. Por outro lado verifica-se que para indicar a sanidade de um indivíduo o teste não é mal pois dos indivíduos negativos para o teste 99,35% são sadios e apenas 0.65% são doentes.

Pensem nesse problema e descubram o falso paradoxo!

Exemplo: As estatísticas da Polícia de uma cidade indicam que 70% dos acidentes de carro ocorrem no fim-de-semana (evento F) e 30% num dia normal (evento N). Um criminalista estudou o gráu alcoólico de motoristas acidentados e obteve os resultados do lado esquerdo da tabela seguinte:

Tabela 8 Distribuições de freqüências do gráu alcoólico por categoria de dia e a Regra de decisão decorrente da fórmula de Bayes

x	P(X=x|N)	P(X=x|F)	P(X=x)	P(N|X=x)	P(F|X=x)
0	0,05	0	0,015	1	0
1	0,10	0	0,030	1	0
2	0.20	0	0.060	1	0
3	0.30	0	0,090	1	0
4	0.20	0,05	0,095	0,630	0,370
5	0,10	0,10	0,100	0,300	0,700
6	0,05	0,20	0,155	0,097	0,903
7	0	0,30	0,210	0	1
8	0	0,20	0,140	0	1
9	0	0,10	0.070	0	1
10	0	0,05	0,035	0	1
	1	1	1		

Usando a fórmula de Bayes, obter uma regra de decisão a respeito do tipo de dia que um acidente ocorreu, conhecido o gráu alcoólico do motorista.

Solução: Primeiramente calculemos P(X=x) que é o denominador da fórmula de Bayes para esse caso. Teremos a aplicação da fórmula da probabilidade total. Observe que a probabilidade total ou também chamada probabilidade marginal corresponde à média ponderada das probabilidades condicionais:

$$P(X=x) = P(F)\cdot P(X=0\mid F) + P(N)\cdot P(X=0\mid N)$$

$$P(X=0) = 0{,}70\cdot 0 + 0{,}30\cdot 0{,}05 = 0{,}015$$

$$P(X=1) = 0{,}70\cdot 0 + 0{,}30\cdot 0{,}10 = 0{,}030$$

$$P(X=2) = 0{,}70\cdot 0 + 0{,}30\cdot 0{,}20 = 0{,}060$$

$$P(X=3) = 0{,}70\cdot 0 + 0{,}30\cdot 0{,}30 = 0{,}090$$

$$P(X=4) = 0{,}70\cdot 0{,}05 + 0{,}30\cdot 0{,}20 = 0{,}095$$

etc.

ficando os resultados completos na quarta coluna da tabela acima. Pela fórmula de Bayes temos:

$$P(N \mid X = k) = \frac{P(N) \cdot P(X = k \mid N)}{P(X = k)} k = 0,1,2,...,10$$

obtendo sucessivamente,

$$P(N \mid X = 0) = \frac{0,3 \cdot 0,05}{0,015} = \frac{0,015}{0,015} = 1$$

$$P(N \mid X = 1) = \frac{0,3 \cdot 0,10}{0,030} = \frac{0,030}{0,030} = 1$$

$$P(N \mid X = 2) = \frac{0,3 \cdot 0,20}{0,060} = \frac{0,060}{0,060} = 1$$

$$P(N \mid X = 3) = \frac{0,3 \cdot 0,30}{0,090} = \frac{0,090}{0,090} = 1$$

$$P(N \mid X = 4) = \frac{0,3 \cdot 0,20}{0,095} = \frac{0,060}{0,095} = 0,630$$

$$P(N \mid X = 5) = \frac{0,3 \cdot 0,10}{0,100} = \frac{0,030}{0,100} = 0,300$$

$$P(N \mid X = 6) = \frac{0,3 \cdot 0,05}{0,155} = \frac{0,030}{0,155} = 0,\ 097 \qquad \text{etc.}$$

completando os resultados na Quinta coluna da tabela acima. Na última coluna da tabela temos a probabilidade do evento complementar F dados os diversos valores de X. Temos então uma regra de decisão para o criminalista: se um motorista acidentado de carro tiver teor alcoólico igual ou superior a sete pode-se concluir com certeza probabilística que o acidente ocorreu num fim-de-semana; caso o teor seja igual ou inferior a três a certeza probabilística é de que o acidente ocorreu num dia normal. Se o resultado do gráu alcoólico estiver entre quatro e seis estamos numa zona de incerteza, ficando a critério do criminalista julgar o dia da semana do acidente com base nas probabilidades "a posteriori" 0,630, 0,300 e 0,097 para X=4, X=5 e X=6 respectivamente. Nesses casos há sempre margem de erro nas decisões, porém conhecidas do criminalista. Esse enfoque de tratamento estatístico constitui o fundamento da Estatística Bayesiana em contraste com a Estatística Clássica.

Completando diremos que os exemplos acima mostram razões da importância da fórmula de Bayes na pesquisa científica. Em áreas em que uma estratégia de decisão é fundamental como em apostas de sucessos de diretrizes econômicas a Estatística Bayesiana é preferível à Estatística Clássica. Neste curso não avançaremos nos princípios bayesianos da Estatística, ficando essa área para cursos mais especializados de mestrado e doutorado.

14. Distribuições Geométrica, Binomial e Binomial Negativa.

Princípios de Contagem

Considere 5 marcas de sabonete e 6 marcas de pasta de dente. Teremos 5·6=30 pares diferentes de (sabonete, pasta de dente). Se a ordem importar então teremos 2·30=60 possibilidades.

Considere as 5 marcas de sabonete. De quantas maneiras posso escolher 3 marcas distintas? Como temos 5 possibilidades de primeira escolha e 4 possibilidades de segunda escolha e 3 de terceira escolha resultam 5·4·3 possibilidades considerando a natureza e a ordem dos ítens como critérios diferenciadores. Enumerando os 60 casos:

123,132,213,231,312,321 **124**,142,214,241,412,421 **125**,152,215,251,512,521
134,143,314,341,413,431 **135**,153,315,351,513,531 **145**,154,415,451,514,541
234,243,324,342,423,432 **235**,253,325,352,523,532 **245**,254,425,452,524,542
345,354,435,453,534,543

Para a enumeração tomamos primeiro os ítens de natureza diferentes que são as **combinações de 5 ítens tomados 3 a 3** e permutamos estes segundo todas as ordens possíveis (**permutações de 3 ítens)** obtendo todos os **arranjos de 5 ítens tomados 3 a 3**. Simbolicamente,

$$A_n^k = C_n^k \cdot P_k$$

Usando a notação fatorial temos no caso particular do exemplo:

$$C_5^3 = \frac{A_5^3}{P_3} = \frac{5 \cdot 4 \cdot 3}{1 \cdot 2 \cdot 3} = \frac{5!}{3!2!}$$

e no caso geral,

$$C_n^k = \frac{A_n^k}{P_k} = \frac{n(n-1)(n-2)...(n-k+1)}{1 \cdot 2 \cdot 3 \cdot ... \cdot k} = \frac{n!}{k!(n-k)!}$$

Uma aplicação importante dessa fórmula é o binômio de Newton escrito assim:

$$(a+b)^n = b^n + C_n^1 ab^{n-1} + C_n^2 a^2 b^{n-2} + C_n^3 a^3 b^{n-3} + ... + C_n^{n-1} a^{n-1} b + a^n$$

Nos casos em que $a = p$ e $b = q$, tais que $p \geq 0$, $q \geq 0$ e $p + q = 1$, teremos:

$$1 = (p+q)^n = q^n + C_n^1 p q^{n-1} + C_n^2 p^2 q^{n-2} + C_n^3 p^3 q^{n-3} + ... + C_n^{n-1} p^{n-1} q + p^n$$

onde temos $n+1$ termos positivos somando 1.

Uma interpretação relevante de combinações poder ser obtida quando queremos obter os modos de colocar três letras A (e duas letras b) em 5 posições . Teremos os 10 casos seguintes:

AAAbb $\Leftrightarrow$	123
AAbAb $\Leftrightarrow$	124
AAbbA $\Leftrightarrow$	125
AbAAb $\Leftrightarrow$	134
AbAbA $\Leftrightarrow$	135
AbbAA $\Leftrightarrow$	145
bAAAb $\Leftrightarrow$	234
bAAbA $\Leftrightarrow$	235
bAbAA $\Leftrightarrow$	245
bbAAA $\Leftrightarrow$	345

correspondendo exatamente às dez combinações obtidas anteriormente pois os números são compostos de dígitos indicadores das posições das letras A's nas cinco posições disponíveis.

Amostra aleatória independente

Ao escrevermos $X_1, X_2, \cdots, X_n$ como uma amostra aleatória independente significamos que repetimos, sob condições idênticas, n experimentos e neles estamos observando uma variável aleatória X . Esses experimentos poderiam ser a jogada de um dado regular e as **n** variáveis definidas pelo número de pontos obtidos em cada uma de **n** jogadas. As distribuições probabilísticas de $X_i, i = 1,2,...,n$ serão idênticas e independentes. Neste caso,

$$P(X_i = k) = \frac{1}{6} \quad para \quad k = 1,2,\cdots,6 \quad e \quad i = 1,2,\cdots,n$$

A independência permite calcular a probabilidade conjunta de eventos pelo produto das probabilidades individuais desses eventos. Por exemplo,

$$P(X_2 = 3; X_4 \neq 6) = P(X_2 = 3) \cdot P(X_4 \neq 6) = \frac{1}{6} \cdot \frac{5}{6} = \frac{5}{36}$$

Uma prova ou experimento chama-se **prova de Bernoulli** caso nessa prova esteja definido apenas um evento **E** de interesse. Numa prova de Bernoulli temos

$$P(E) = p \text{ e } P(\overline{E}) = 1 - p = q$$

Uma seqüência de **n** provas idênticas e independentes de Bernoulli pode ser escrita da seguinte forma:

$$E_1, E_2, \cdots, E_n \xrightarrow{IID} P$$

Vejamos alguns exemplos para fixar esses conceitos:

Exemplo (Distribuição Geométrica) Obtenha a distribuição do número de tentativas N_1 necessárias para sair o primeiro número 6 numa amostra aleatória independente com tamanho indefinido.

Solução: Chamemos de 6 e $\bar{6}$ a ocorrência e a não ocorrência do 6 numa tentativa. Temos então que $P(6) = \frac{1}{6}$ e $P(\overline{6}) = \frac{5}{6}$, donde pela independência das tentativas,

$$P(N_1 = 1) = P(6) = \frac{1}{6} \qquad P(N_1 = 2) = P(\bar{6} \cap 6) = \frac{5}{6} \cdot \frac{1}{6} \qquad P(N_1 = 3) = P(\bar{6} \cap \bar{6} \cap 6) = \frac{5}{6} \cdot \frac{5}{6} \cdot \frac{1}{6}$$

Generalizando tem-se pois $$P(N_1 = k) = \left(\frac{5}{6}\right)^{k-1} \cdot \frac{1}{6} \qquad k = 1,\ 2,\ 3,\ \cdots$$

Essa distribuição é chamada de **distribuição geométrica**. Decorre esse nome de uma demonstração probabilística para o limite da soma parcial uma progressão geométrica, pois
como a soma das probabilidades deve ser 1, obtemos

$$1 = \frac{1}{6} \cdot \left(1 + \frac{5}{6} + \left(\frac{5}{6}\right)^2 + \left(\frac{5}{6}\right)^3 + \cdots\right)$$

donde,

$$1 + \frac{5}{6} + \left(\frac{5}{6}\right)^2 + \left(\frac{5}{6}\right)^3 + \cdots = 6 = \frac{1}{1 - \frac{5}{6}}$$

ou seja,

$$\lim_{n\to\infty}\sum_{k=0}^{n}\left(\frac{5}{6}\right)^{k}=\frac{1}{1-\frac{5}{6}}$$

Se chamarmos de **p** a probabilidade de falha e de **q** a probabilidade de sucesso de um equipamento num dia, então N_1 pode ser interpretada como o número de dias decorridos até que a primeira falha do equipamento ocorra. Seguindo os mesmos passos chegaremos à distribuição geométrica dessa variável aleatória:

$$P(N_1 = n_1) = q^{n_1 - 1} \cdot p \quad n_1 = 1,2,\cdots$$

e à demonstração de que

$$1+q+q^2+q^3+\cdots=\frac{1}{1-q}$$

$$\lim_{n\to\infty}\sum_{i=0}^{n}q^{i}=\frac{1}{1-q}$$

de forma probabilística.

Exemplo (Distribuição Binomial) Numa amostra aleatória independente observam-se **n** provas de Bernoulli onde $P(E)=p$ e $P(\overline{E})=1-p=q$. Obtenha a distribuição do número **K** de ocorrências de E. Na linguagem de eventos podemos escrever:

$$\Omega=(E_1+\overline{E}_1)\cap(E_2+\overline{E}_2)\cap(E_3+\overline{E}_3)\cap\cdots(E_n+\overline{E}_n)$$

isto é, com certeza o evento ocorre só na primeira vez, ocorre só na segunda vez, ocorre só na terceira vez, ocorre apenas na primeira e na segunda vez, etc. Distribuindo essas interseções teremos:

$$\begin{aligned}\Omega = \quad & E_1\cap E_2\cap\cdots\cap E_{n-1}\cap E_n+\\ & \overline{E}_1\cap E_2\cap\cdots\cap E_{n-1}\cap E_n+\\ & \cdots\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots\ +\\ & E_1\cap E_2\cap\cdots\cap E_{n-1}\cap\overline{E}_n+\\ & \overline{E}_1\cap\overline{E}_2\cap\cdots\cap E_{n-1}+E_n+\\ & \cdots\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots\ +\\ & E_1\cap E_2\cap\cdots\cap\overline{E}_{n-1}\cap\overline{E}_n+\\ & \cdots\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots\ +\\ & \overline{E}_1\cap\overline{E}_2\cap\cdots\cap\overline{E}_{n-1}\cap\overline{E}_n\end{aligned}$$

isto é, interpretando na álgebra de eventos, $\Omega=(\overline{E}+E)^n$, resultando:

$$P(\Omega) = p^n + C_n^1 q p^{n-1} + C_n^2 q^2 p^{n-2} + C_n^3 q^3 p^{n-3} + \ldots + C_n^{n-1} q^{n-1} p + q^n = 1$$

ou seja, será 1 a soma da probabilidade de ocorrerem **n** vezes o evento **E** mais a probabilidade de ocorrerem **n-1** vezes mais a probabilidade de ocorrerem **n-2** vezes até a última parcela que será a probabilidade de não ocorrência de **E** em nenhuma das **n** vezes. Desse raciocínio decorre então a **distribuição binomial** de probabilidades:

$$P(K = k) = C_n^k p^k q^{n-k} k = 0,1,2,\cdots,n$$

Exemplo. (Distribuição Binomial Negativa) Seja { $E_1, E_2, \cdots, E_n, \cdots$ } uma seqüência de provas idênticas e independentes de Bernoulli com probabilidade **p** de ocorrer o evento E. Vamos obter as distribuições de **N_2, N_3, ... , N_v,** os números necessários de provas para que ocorra o evento E pela 2ª, 3ª, ... ,vª vez, respectivamente.

Comecemos por N_2: os possíveis valores de N_2 serão 2, 3, 4, ... , k, ...O espaço amostral dessas provas induzido por N_2 e as probabilidades serão dadas pela tabela seguinte:

n_2	ω	P(ω)	$P(N_2=n_2)$
2	EE	p^2	p^2
3	$E\overline{E}E$	p^2q	
3	$\overline{E}EE$	p^2q	$2p^2q$
4	$\overline{E}\overline{E}EE$	p^2q^2	
4	$\overline{E}E\overline{E}E$	p^2q^2	
4	$E\overline{E}\overline{E}E$	p^2q^2	$3p^2q^2$
5	$E\overline{E}\overline{E}\overline{E}E$	p^2q^3	
5	$\overline{E}E\overline{E}\overline{E}E$	p^2q^3	
5	$\overline{E}\overline{E}E\overline{E}E$	p^2q^3	
5	$\overline{E}\overline{E}\overline{E}EE$	p^2q^3	$4p^2q^3$
...	...	...	...

Da construção acima resulta que

$$P(N_2 = n_2) = C_{k-1}^1 p^2 q^{n_2-2}, para n_2 = 2,3,4,5,...$$

Avancemos para N_3: os possíveis valores de N_3 serão 3, 4, 5, ..., k, ... A tabela seguinte mostra a distribuição de N_3:

k	ω	P(ω)	P(N$_3$=k)
3	EEE	p^3	$C_2^0 p^3$
4	$EE\overline{E}E$	qp^3	
4	$E\overline{E}EE$	qp^3	
4	$\overline{E}EEE$	qp^3	$C_3^1 p^3$
5	$\overline{E}\overline{E}EEE$	$q^2 p^3$	
5	$\overline{E}E\overline{E}EE$	$q^2 p^3$	
5	$\overline{E}EE\overline{E}E$	$q^2 p^3$	
5	$E\overline{E}\overline{E}EE$	$q^2 p^3$	
5	$E\overline{E}E\overline{E}E$	$q^2 p^3$	
5	$EE\overline{E}\overline{E}E$	$q^2 p^3$	$C_4^2 q^2 p^3$
..	...	...	...

Obtemos nesse caso a expressão $P(N_3 = k) = C_{k-1}^{k-3} q^{k-3} p^3$

Para o caso geral N_v vamos, pelo mesmo caminho, obter a fórmula geral da distribuição binomial negativa:

$$P(N_v = k) = C_{k-1}^{k-v} q^{k-v} p^v, para k = v, v+1, v+2, \cdots$$

ou

$$P(N_v = k) = C_{k-1}^{v-1} q^{k-v} p^v, para k = v, v+1, v+2, \cdots$$

De forma probabilística obtemos que

$$1 = \sum_{k=v}^{\infty} C_{k-1}^{k-v} q^{k-v} p^v$$

$$1 = p^v + C_v^1 q \; p^v + C_{v+1}^2 q^2 p^v + C_{v+2}^3 q^3 p^v + \cdots$$

Fatorando p^v, lembrando que $p = 1 - q$ e passando para o lado esquerdo, temos

$$\frac{1}{(1-q)^v} = 1 + C_v^1 q \; + C_{v+1}^2 q^2 + C_{v+2}^3 q^3 + \cdots$$

$$(1-q)^{-v} = 1 + C_v^1 q \; + C_{v+1}^2 q^2 + C_{v+2}^3 q^3 + \cdots$$

que representa a extensão do binômio de Newton para expoente negativo, justificando a denominação da distribuição de N_v.

15. Problemas para Revisão

Problema 1 Considere a jogada de dois dados regulares. Obtenha a distribuição conjunta de $U = Max(X, Y)$ e $V = Min(X, Y)$ e as distribuições marginais decorrentes. Após isso calcule $P(V \geq 3; U \leq 5)$, $P(V \geq 3)$ e $P(U \leq 5)$.

Resposta

A distribuição conjunta de U e V

u \ v	1	2	3	4	5	6	P(u)
1	1/36	0	0	0	0	0	**1/36**
2	2/36	1/36	0	0	0	0	**3/36**
3	2/36	2/36	1/36	0	0	0	**5/36**
4	2/36	2/36	2/36	1/36	0	0	**7/36**
5	2/36	2/36	2/36	2/36	1/36	0	**9/36**
6	2/36	2/36	2/36	2/36	2/36	1/36	**11/36**
P(v)	**11/36**	**9/36**	**7/36**	**5/36**	**3/36**	**1/36**	**1**

$P(V \geq 3; U \leq 5) = \frac{1}{4}$ $\quad P(V \geq 3) = \frac{4}{9}$ $\quad P(U \leq 5) = \frac{25}{36}$

Problema 2 A distribuição conjunta de X e Y está apresentada na seguinte tabela:

x \ y	2	3	4	6	9	$P(x)$
1	1/24	1/24	2/24	1/24	1/24	**6/24**
2	1/24	2/24	3/24	2/24	1/24	**9/24**
3	0	1/24	2/24	1/24	0	**4/24**
4	1/24	1/24	1/24	1/24	1/24	**5/24**
$P(y)$	**3/24**	**5/24**	**8/24**	**5/24**	**3/24**	**24/24**

a) Obtenha as distribuições da soma $S = X + Y$ e do módulo da diferença $D = |X - Y|$, separadamente. Faça o histograma dessas distribuições. Calcule $P(S \leq 4)$ e $P(D \geq 3)$.

b) Calcule os dois conjuntos de distribuições condicionais $P_x(y)$ e $P_y(x)$.

c) Com base no ítem b) discuta a independência ou não das variáveis aleatórias X e Y.

Resposta

Distribuição da soma S=X+Y

s	3	4	5	6	7	8	9	10	11	12	13	**Total**
P(s)	1/24	2/24	4/24	5/24	4/24	3/24	1/24	2/24	1/24	0	1/24	1

Distribuição de D=\|X-Y\|

D	0	1	2	3	4	5	6	7	8	**Total**
P(d)	3/24	6/24	6/24	3/24	2/24	2/24	0	1/24	1/24	1

$P(S \leq 4) = 1/8$ $P(D \geq 3) = 3/8$

Distribuições condicionais $P_x(y)$

y / X	**2**	**3**	**4**	**6**	**9**	**Total**
$P_1(y)$	1/6	1/6	2/6	1/6	1/6	**1**
$P_2(y)$	1/9	2/9	3/9	2/9	1/9	**1**
$P_3(y)$	0	1/4	2/4	1/4	0	**1**
$P_4(y)$	1/5	1/5	1/5	1/5	1/5	**1**
P(y)	**3/24**	**5/24**	**8/24**	**5/24**	**3/24**	**1**

Distribuições condicionais $P_y(x)$

y / x	**$P_2(x)$**	**$P_3(x)$**	**$P_4(x)$**	**$P_6(x)$**	**$P_9(x)$**	$P(x)$
1	1/3	1/5	2/8	1/5	1/3	**6/24**
2	1/3	2/5	3/8	2/5	1/3	**9/24**
3	0	1/5	2/8	1/5	0	**4/24**
4	1/3	1/5	1/8	1/5	1/3	**5/24**
Total	**1**	**1**	**1**	**1**	**1**	**1**

Como as condicionais não são iguais como função, X e Y são dependentes

Problema 3 Considere a variável aleatória T_1, o tempo gasto pelo trabalhador 1 para realizar uma certa tarefa, com a seguinte distribuição:

$$P(T_1 = 2k+3) = C_3^k \left[\frac{1}{2}\right]^3, \ para\ k = 0,1,2\ e\ 3$$

a) Calcule $P(T_1 \leq 7)$, $P(T_1 \geq 5)$ e $P(|T_1 - 6| \leq 1)$.

b) Dois trabalhadores 1 e 2 são convocados para realizar duas tarefas seqüencialmente. Qual a distribuição do tempo total gasto pelos dois ($T = T_1 + T_2$) ? (Suponha que haja independência entre os tempos de cada um).

Resposta

Distribuições Marginais (pela fórmula) e **Conjunta** (pela independência)

T_1 \ t_2	3	5	7	9	$P(t_1)$
3	1/64	3/64	3/64	1/64	**1/8**
5	3/64	9/64	9/64	3/64	**3/8**
7	3/64	9/64	9/64	3/64	**3/8**
9	1/64	3/64	3/64	1/64	**1/8**
$P(t_2)$	**1/8**	**3/8**	**3/8**	**1/8**	**1**

$$P(T_1 \leq 7) = \frac{7}{8} \qquad P(T_1 \geq 5) = \frac{7}{8} \qquad P(|T_1 - 6| \leq 1) = \frac{3}{4}$$

Valores de $T=T_1+T_2$ (Tabela auxiliar)

T_1 \ t_2	3	5	7	9
3	6	8	10	12
5	8	10	12	14
7	10	12	14	16
9	12	14	16	18

t	**6**	**8**	**10**	**12**	**14**	**16**	**18**	**Total**
P(t)	1/64	6/64	15/64	20/64	15/64	6/64	1/64	**1**

Problema 4 Uma turma tem 6 mulheres e 4 homens. Seis pessoas são sorteadas para uma comissão. Obtenha a distribuição de probabilidades do número **K** de mulheres presentes numa comissão. Após, obtenha $P(K \leq 5)$.

Resposta: Como é 6 o tamanho da comissão, o menor número de homens nela será 2. O espaço amostral será $\Omega = \{2,3,4,5,6)$. Como não pode haver repetição de membro da comissão teremos:

Distribuição do número **K** de mulheres na comissão

k	P(K=k)
2	$P(K=2) = \frac{C_6^2 \cdot C_4^4}{C_{10}^6} = \frac{15}{210}$
3	$P(K=3) = \frac{C_6^3 \cdot C_4^3}{C_{10}^6} = \frac{80}{210}$
4	$P(K=4) = \frac{C_6^4 \cdot C_4^2}{C_{10}^6} = \frac{90}{210}$
5	$P(K=5) = \frac{C_6^5 \cdot C_4^1}{C_{10}^6} = \frac{24}{210}$
6	$P(K=6) = \frac{C_6^6 \cdot C_4^0}{C_{10}^6} = \frac{1}{210}$
Total	1

Uma distribuição desse tipo é chamada **Distribuição Hipergeométrica**.

Problema 5 Sob nenhum conhecimento **a priori** a respeito de qual a resposta produtiva à adubação nitrogenada de uma determinada variedade de soja um pesquisador da EMBRAPA obteve as distribuições da produção em kg/parcela em função do nível de nitrogênio aplicado constantes na tabela seguinte. Obtenha uma regra de decisão estatística que indique a um cientista do solo qual o nível de adubação nitrogenada sendo utilizada numa propriedade, conhecida sua produção, ou seja, obtenha P(N=0|x), P(N=1|x), P(N=2|x) e P(N=3|x) para x=20, 22, ... , 38 seguindo o modelo do criminalista em exemplo dado.

Produção Média (**x**)	P(x|N=0)	P(x|N=1)	P(x|N=2)	P(x|N=3)
20	0,1	0	0	0
22	0,4	0,1	0	0
26	0,4	0,2	0	0
28	0,1	0,4	0,1	0
30	0	0.2	0,3	0
32	0	0.1	0,3	0,1
34	0	0	0,2	0,4
36	0	0	0,1	0,4
38	0	0	0	0,1

Resposta:

Produção Média(**x**)	P(x|N=0)	P(x|N=1)	P(x|N=2)	P(x|N=3)	P(x)	$P_x(N=0)$	$P_x(N=1)$	$P_x(N=2)$	$P_x(N=3)$
20	0,1	0	0	0	0,025	1	0	0	0
22	0,4	0,1	0	0	0,125	0,8	0.2	0	0
26	0,4	0,2	0	0	0,150	0,67	0,33	0	0
28	0,1	0,4	0,1	0	0,150	0,165	0,670	0,165	0
30	0	0.2	0,3	0	0,125	0	0,4	0.6	0
32	0	0.1	0,3	0,1	0,125	0	0,2	0.6	0.2
34	0	0	0,2	0,4	0,150	0	0	0.33	0.67
36	0	0	0,1	0,4	0,125	0	0	0.2	0.8
38	0	0	0	0,1	0,025	0	0	0	1
P(n)	1/4	1/4	1/4	1/4	1	1/4	1/4	1/4	1/4

Problema 6 Seja X uma variável aleatória assumindo valores inteiros x_1, x_2, $x_3, \ldots$ com distribuição P(x). A **esperança matemática de h(X)** , simbolizada por E[h(X)], será definida por

$$E[h(X)] = h(x_1)P(x_1) + h(x_2)P(x_2) + \cdots + h(x_k)P(x_k) + \cdots$$

Sejam X e Y os pontos de dois dados e U=Max(X,Y). Usando a definição de esperança matemática obtenha:

a) $\mu_U = E(U)$ b) $E(U^2)$

c) $\sigma_U{}^2 = E[(U - \mu_U)^2]$ d) Verifique que $\sigma_U^2 = E(U^2) - \mu_U^2$

Resposta: No problema 1 obtemos a tabela:

u \ v	1	2	3	4	5	6	P(u)
1	1/36	0	0	0	0	0	**1/36**
2	2/36	1/36	0	0	0	0	**3/36**
3	2/36	2/36	1/36	0	0	0	**5/36**
4	2/36	2/36	2/36	1/36	0	0	**7/36**
5	2/36	2/36	2/36	2/36	1/36	0	**9/36**
6	2/36	2/36	2/36	2/36	2/36	1/36	**11/36**
P(v)	**11/36**	**9/36**	**7/36**	**5/36**	**3/36**	**1/36**	**1**

onde P(u) é a distribuição de U. Portanto, pelas definições:

$$\mu_U = E(U) = 1\cdot\frac{1}{36} + 2\cdot\frac{3}{36} + 3\cdot\frac{5}{36} + 4\cdot\frac{7}{36} + 5\cdot\frac{9}{36} + 6\cdot\frac{11}{36} = \frac{1+6+15+28+45+66}{36} = \frac{161}{36}$$

$$E(U^2) = 1^2\cdot\frac{1}{36} + 2^2\cdot\frac{3}{36} + 3^2\cdot\frac{5}{36} + 4^2\cdot\frac{7}{36} + 5^2\cdot\frac{9}{36} + 6^2\cdot\frac{11}{36} = \frac{1+12+45+112+225+396}{36} = \frac{791}{36}$$

$$\sigma_U{}^2 = E(U-\mu_U)^2 =$$

$$\left(1-\frac{161}{36}\right)^2\frac{1}{36}+\left(2-\frac{161}{36}\right)^2\frac{3}{36}+\left(3-\frac{161}{36}\right)^2\frac{5}{36}+\left(4-\frac{161}{36}\right)^2\frac{7}{36}+\left(5-\frac{161}{36}\right)^2\frac{9}{36}+\left(6-\frac{161}{36}\right)^2\frac{11}{36}=$$

$$\frac{15625+23763+14045+2023+3249+33275}{36^3} = \frac{91980}{36^3} = \frac{2555}{36^2}$$

Mas, $$\sigma_U^2 = E(U^2) - \mu_U^2 = \frac{791}{36} - \left(\frac{161}{36}\right)^2 = \frac{28476 - 25921}{36^2} = \frac{2555}{36^2}$$

demonstrando a igualdade indicada.

Problema 7 a) Obtenha a distribuição do número de partos necessários para o nascimento do primeiro menino (N_1). Qual a probabilidade do primeiro menino nascer após o terceiro parto? b) Obtenha a distribuição do número necessário de jogadas de uma moeda para a saída da Segunda "cara". Qual a probabilidade da Segunda "cara" acontecer antes da quarta jogada?

Resposta: Da parte a) temos a distribuição geométrica (ou binomial negativa com v=1).

$$P(N_1 = n_1) = q^{n_1-1}p, para n_1 = 1,2,3,\cdots$$

onde $p = \frac{1}{2}$ e $q = \frac{1}{2}$, resultando na distribuição:

n_1	$P(n_1)$
1	1/2
2	1/4
3	1/8
4	1/16
...	...

Mas, $$P(N_1 > 3) = 1 - P(N_1 \leq 3) = 1 - \left(\frac{1}{2} + \frac{1}{4} + \frac{1}{8}\right) = 1 - \frac{7}{8} = \frac{1}{8}$$

Na parte b) temos a binomial negativa com $v = 2 \ e \ p = \dfrac{1}{2}$, donde

$$P(N_2 = n_2) = C^1_{n_2 - 1}(\frac{1}{2})^{n_2 - 2}(\frac{1}{2})^2, \ para \ n_2 = 2,3,4,\cdots$$

$$P(N_2 = n_2) = (n_2 - 1)(\frac{1}{2})^{n_2 - 2}\frac{1}{4}, \ para \ n_2 = 2,3,4,\cdots$$

$$P(N_2 = n_2) = (n_2 - 1)(\frac{1}{2})^{n_2}, \ para \ n_2 = 2,3,4,\cdots$$

n_2	$P(n_2)$
2	1/4
3	2/8
4	3/16
5	4/32
...	...

donde, $P(N_2 < 4) = \dfrac{1}{4} + \dfrac{2}{8} = \dfrac{1}{2}$.

Problema 8 De dez peixes de uma represa cinco são marcados. Pescam-se quatro peixes **com reposição** dos peixes no reservatório. Qual a distribuição de probabilidades do número **K** de peixes marcados pescados? Qual a probabilidade de todos serem marcados e a de que nenhum seja marcado?

Resposta

Aqui temos o modelo da distribuição binomial com **n=**4 e **p=**5/10=1/2 e a fórmula

$$P(K = k) = C^k_4\left(\frac{1}{2}\right)^k\left(\frac{1}{2}\right)^{4-k} = C^k_4\left(\frac{1}{2}\right)^4 = \frac{C^k_4}{16}k = 0,1,\cdots,4$$

$P(K = 4) = \dfrac{C^4_4}{16} = \dfrac{1}{16}$ (todos) $\qquad P(K = 0) = \dfrac{C^0_4}{16} = \dfrac{1}{16}$ (nenhum)

Problema 9 Supondo iguais as probabilidades de nascer menino ou menina, obtenha a probabilidade do terceiro menino nascer após o quinto parto.

Resposta: Aqui é um caso de binomial negativa com **v=**3 e **p=**1/2.

$$P(N_3 = n_3) = C^2_{n_3-1}(\frac{1}{2})^{n_3-3}(\frac{1}{2})^3,\ para\ n_3 = 3,4,5,\cdots$$

$$P(N_3 = n_3) = \frac{(n_3-1)(n_3-2)}{2}(\frac{1}{2})^{n_3},para\ n_3 = 3,4,5,\cdots$$

n_3	$P(n_3)$
3	1/8
4	3/16
5	6/32
6	10/64
...	...

$$P(N_3 > 5) = 1 - P(N_3 \le 5) = 1 - \left(\frac{1}{8} + \frac{3}{16} + \frac{6}{32}\right) = 1 - \frac{4+6+6}{32} = 1 - \frac{16}{32} = \frac{16}{32} = \frac{1}{2}$$

Problema 10 Suponhamos que a probabilidade do Bahia ganhar do Vitória seja de 0,4. Qual a probabilidade do Bahia ganhar pela segunda vez do Vitória até a quinta partida consecutiva?

Resposta: Como no problema 7 aplicamos a binomial negativa para v=2 e p=1/4.

$$P(N_2 = n_2) = C^1_{n_2-1}(\frac{6}{10})^{n_2-2}(\frac{4}{10})^2,paran_2 = 2,3,4,\cdots$$

$$P(N_2 = n_2) = (n_2-1)(\frac{6}{10})^{n_2-2}(\frac{4}{10})^2,paran_2 = 2,3,4,\cdots$$

n_2	$P(n_2)$
2	16000/100000
3	19200/100000
4	17280/100000
5	13824/100000
...	...

$$P(N_2 \le 5) = \frac{66304}{100000} = 0{,}66304$$

Problema 11 Sejam 5 bolas numeradas com os respectivos pesos: 1, 2, 3, 2, 1. Duas bolas são tomadas simultaneamente e são pesadas. Obter a distribuição do peso **V** das bolas retiradas. Calcule $P(V \ge 4)$ e $E(V^2)$.

Resposta: O espaço probabilística será:

ω	$P(\omega)$	$v = V(\omega)$
(1,2)	1/10	3
(1,3)	1/10	4
(1,4)	1/10	3
(1,5)	1/10	2
(2,3)	1/10	5
(2,4)	1/10	4
(2,5)	1/10	3
(3,4)	1/10	5
(3,5)	1/10	4
(4,5)	1/10	3
	1	

Decorre imediatamente que

v	2	3	4	5	**Total**
P(v)	1/10	4/10	3/10	2/10	**1**

$$P(V \ge 4) = \frac{5}{10} = \frac{1}{2} \quad E(V^2) = 2^2\frac{1}{10} + 3^2\frac{4}{10} + 4^2\frac{3}{10} + 5^2\frac{2}{10} = \frac{4+36+48+50}{10} = \frac{138}{10} = 13{,}8$$

Problema 12 A tabela abaixo mostra as distribuições condicionais de X dado Y e a marginal de Y. a) Obtenha a distribuição marginal de X; b) Obtenha a distribuição conjunta de X e Y; c) Calcule a distribuição de S=X+Y e D=|X-Y|; d) Calcule E(S) e E(D).

y x	1	2	3	P(x)
0	1/8	1/8	1/8	**( )**
1	2/8	2/8	2/8	**( )**
2	3/8	3/8	3/8	**( )**
3	2/8	2/8	2/8	**( )**
Total	**1**	**1**	**1**	<><><>
P(y)	**1/3**	**1/3**	**1/3**	1

Resposta:

Distribuição conjunta de X e Y e a marginal de X

y X	1	2	3	P(x)
0	1/24	1/24	1/24	**1/8**
1	2/24	2/24	2/24	**2/8**
2	3/24	3/24	3/24	**3/8**
3	2/24	2/24	2/24	**2/8**
P(y)	**1/3**	**1/3**	**1/3**	1

Tabela auxiliar para S

y x	1	2	3
0	1	2	3
1	2	3	4
2	3	4	5
3	4	5	6

Distribuição de S

s	1	2	3	4	5	6	Total
P(s)	1/24	3/24	6/24	7/24	5/24	2/24	**1**

Tabela auxiliar para D

x \ Y	1	2	3
0	1	2	3
1	0	1	2
2	1	0	1
3	2	1	0

Distribuição de D

D	**0**	**1**	**2**	**3**	**Total**
P(d)	7/24	11/24	5/24	1/24	**1**

$$E(S) = 1\frac{1}{24} + 2\frac{3}{24} + 3\frac{6}{24} + 4\frac{7}{24} + 5\frac{5}{24} + 6\frac{2}{24} = \frac{1+6+18+28+25+12}{24} = \frac{90}{24} = 3{,}75$$

$$E(D) = 0\frac{7}{24} + 1\frac{11}{24} + 2\frac{5}{24} + 3\frac{1}{24} = \frac{0+11+10+3}{24} = \frac{24}{24} = 1$$

Problema 13: O componente 1 tem probabilidade de falha de 0,1 e o componente 2, de 0,2. Sejam

X: Número de testes para o componente 1 apresentar a primeira falha

Y: Número de testes para o componente 2 apresentar a primeira falha

a) Complete as tabelas abaixo, supondo X e Y independentes

x	P(x)
1	
2	
3	
4	
5	
>5	
	1

y	P(y)
1	
2	
3	
4	
5	
>5	
	1

x \ y	1	2	3	4	5	>5	P(x)
1							
2							
3							
4							
5							
>5							
P(y)							1

b) Calcule $P(X > 3)$; $P(Y \leq 4)$; $P(X < 3; Y < 4)$ e $P(X > 5; Y > 4)$

c) Calcule a distribuição (censurada) conjunta de X e Y condicionada à ocorrência do evento $\{X \leq 5; Y \leq 5\}$ e recalcule as probabilidades da parte b).

Resposta: A distribuição da primeira falha é geométrica:

$$P(N_1 = n_1) = q^{n_1 - 1} p \text{ para } n_1 = 1, 2, \ldots$$

Substituindo p = 0,1 e p = 0,2 para X e Y respectivamente, teremos:

x	P(x)
1	0,10
2	0,09
3	0,081
4	0,0729
5	0,0656
>5	0,5905
	1

y	P(y)
1	0,20
2	0,16
3	0,128
4	0,1024
5	0,0819
>5	0,3277
	1

x \ y	1	2	3	4	5	>5	P(x)
1	0,0200	0,0160	0,0128	0,0102	0,0082	0,0328	**0,10**
2	0,0180	0,0144	0,0115	0,0092	0,0074	0,0295	**0,09**
3	0,0162	0,0130	0,0104	0,0083	0,0066	0,0265	**0,081**
4	0,0146	0,0117	0,0093	0,0075	0,0059	0,0239	**0,0729**
5	0,0131	0,0105	0,0084	0,0067	0,0054	0,0215	**0,0656**
>5	0,1181	0,0944	0,0756	0,0605	0,0484	0,1935	**0,5905**
P(y)	**0,2**	**0,16**	**0,128**	**0,1024**	**0,0819**	**0,3277**	**1**

$P(X > 3) = 0{,}729$ $\quad$ $P(Y \leq 4) = 0{,}5904$ $\quad$ $P(X < 3; Y < 4) = 0{,}0927$ $\quad$ $P(X > 5; Y > 4) = 0{,}2419$

Distribuição censurada

x \ y	1	2	3	4	5	P(x)
1	0,0726	0,0581	0,0465	0,0370	0,0298	**0,2440**
2	0,0654	0,0524	0,0418	0,0334	0,0269	**0,2199**
3	0,0588	0,0472	0,0379	0,0301	0,0240	**0,1980**
4	0,0531	0,0425	0,0338	0,0272	0,0214	**0,1780**
5	0,0476	0,0381	0,0305	0,0243	0,0196	**0,1601**
P(y)	**0,2975**	**0,2383**	**0,1905**	**0,1520**	**0,1217**	**1**

Problema 14: Uma urna contém 9 bolas sendo 2 brancas, 3 pretas e 4 vermelhas. Três bolas são retiradas de uma vez e

X = Número de bolas brancas e

Y = Número de bolas pretas

são observadas.

a) Usando a fórmula de combinação de **n** objetos tomados **k a k**, obtenha:

$P(X = x; Y = y)$ para x = 0, 1 e 2 e y = 0, 1, 2 e 3 preenchendo a tabela da distribuiçãoconjunta de X e Y.

b) Obtenha as tabelas de P(x|y) e P(y|x)

d) Discuta independência de X e Y

e) Calcule E(X) e E(Y)

Resposta:

x \ y	0	1	2	3	P(x)
0	$\frac{C_4^3}{C_9^3}=\frac{4}{84}$	$\frac{C_3^1C_4^2}{C_9^3}=\frac{18}{84}$	$\frac{C_3^2C_4^1}{C_9^3}=\frac{12}{84}$	$\frac{C_3^3}{C_9^3}=\frac{1}{84}$	$\frac{35}{84}$
1	$\frac{C_2^1C_4^2}{C_9^3}=\frac{12}{84}$	$\frac{C_2^1C_3^1C_4^1}{C_9^3}=\frac{24}{84}$	$\frac{C_2^1C_3^2}{C_9^3}=\frac{6}{84}$	$\frac{0}{84}$	$\frac{42}{84}$
2	$\frac{C_2^2C_4^1}{C_9^3}=\frac{4}{84}$	$\frac{C_2^2C_3^1}{C_9^3}=\frac{3}{84}$	$\frac{0}{84}$	$\frac{0}{84}$	$\frac{7}{84}$
P(y)	$\frac{20}{84}$	$\frac{45}{84}$	$\frac{18}{84}$	$\frac{1}{84}$	$\frac{84}{84}=1$

x \ y	0	1	2	3	P(x)
0	1/5	6/15	2/3	1	**5/12**
1	3/5	8/15	1/3	0	**6/12**
2	1/5	1/15	0	0	**1/12**
Total	**1**	**1**	**1**	**1**	**1**

x \ y	0	1	2	3	Total
0	4/35	18/35	12/35	1/35	**1**
1	2/7	4/7	1/7	0	**1**
2	4/7	3/7	0	0	**1**
P(y)	**20/84**	**45/84**	**18/84**	**1/84**	**1**

Como as distribuições condicionais não são iguais entre si, X e Y são dependentes

$$E(X) = 0 \cdot \frac{35}{84} + 1 \cdot \frac{42}{84} + 2 \cdot \frac{7}{84} = \frac{56}{84} = \frac{2}{3} \qquad E(Y) = 0 \cdot \frac{20}{84} + 1 \cdot \frac{45}{84} + 2 \cdot \frac{18}{84} + 3 \cdot \frac{1}{84} = \frac{84}{84} = 1$$

Problema 15: Com dados da CEF sabe-se que P(Y=1) = 0,6 ; P(Y=2) = 0,2 e P(Y=3) = 0,2, onde Y é o resultado "coluna 1", "coluna 2" e "coluna 3". Após uma pesquisa de jogos anteriores acompanhados dos palpites de conhecedores de futebol ele obteve a parte esquerda da tabela abaixo, onde X é o palpite de um conhecedor de futebol.

x	P(x |y= 1)	P(x |y= 2)	P(x |y= 3)	P(x)	P(y =1 |x)	P(y =2 |x)	P(y =3 |x)
1	0,8	0,2	0,1				
2	0,1	0,6	0,2				
3	0,1	0,2	0,7				
Total	1	1	1	1			
P(y)	( )	( )	( )	1	( )	( )	( )

Obtenha regras de decisão para apontar a coluna dado o palpite do conhecedor.

Resposta:

x	**P(x |y= 1)**	**P(x |y= 2)**	**P(x |y= 3)**	**P(x)**	**P(y =1 |x)**	**P(y =2 |x)**	**P(y =3 |x)**	**Total**
1	0,8	0,2	0,1	0,54	0,88889	0,07407	0,03704	1
2	0,1	0,6	0,2	0,22	0,27273	0,54545	0,18182	1
3	0,1	0,2	0,7	0,24	0,25000	0,16667	0,58333	1
Total	1	1	1	1				
P(y)	(0,6)	(0,2)	(0,2)	1	(0,6)	(0,2)	(0,2)	1

Problema 16: Considere a tabela à esquerda das distribuições condicionais de Y dados X e a distribuição marginal de X.

Obtenha a distribuição conjunta de X e Y e a marginal de Y, completando a tabela à direita.

a) Seja S = X+Y. Calcule a distribuição de S e E(S)
b) Seja V = Min(X,Y). Calcule a distribuição de V e E(V)

	y	2	4	6	**Total**	**P(x)**
x						
1		1/4	2/4	1/4	**1**	**0,5**
3		3/8	3/8	2/8	**1**	**0,3**
5		0	1/2	1/2	**1**	**0,2**
P(y)					**1**	**1**

	y	2	4	6	**P(x)**
x					
1		(10/80)	(20/80)	(10/80)	**(40/80)**
3		(9/80)	(9/80)	(6/80)	**(24/80)**
5		(0/80)	(8/80)	(8/80)	**(16/80)**
P(y)		**(19/80)**	**(37/80)**	**(24/80)**	**1**

Resposta:

Valores de S = X+Y

	y	2	4	6
x				
1		3	5	7
3		5	7	9
5		7	9	11

Valores de V = Min(X,Y)

	y	2	4	6
x				
1		1	1	1
3		2	3	3
5		2	4	5

Distribuição de S

s	**3**	**5**	**7**	**9**	**11**	**Total**
P(s)	10/80	29/80	19/80	14/80	8/80	**1**

$$E(S) = 3 \cdot \frac{10}{80} + 5 \cdot \frac{29}{80} + 7 \cdot \frac{19}{80} + 9 \cdot \frac{14}{80} + 11 \cdot \frac{8}{80} = \frac{522}{80} = 6{,}525$$

Distribuição de V

v	**1**	**2**	**3**	**4**	**5**	**Total**
P(v)	40/80	9/80	15/80	8/80	8/80	**1**

$$E(V) = 1 \cdot \frac{40}{80} + 2 \cdot \frac{9}{80} + 3 \cdot \frac{15}{80} + 4 \cdot \frac{8}{80} + 5 \cdot \frac{8}{80} = \frac{175}{80} = \ 2{,}1875$$

Parâmetros de uma Distribuição de Probabilidades

Sejam X e Y duas variáveis aleatórias com distribuição conjunta dada por

$$P(X = x_i ; Y = y_j) = p_{ij} (i = 1,2,3,... j = 1,2,3,...)$$

Caso os espaços amostrais sejam finitos então **m** e **n** serão os limites superiores para **i** e **j** respectivamente. Enumerando as probabilidades na **notação simplificada**, onde o ponto significa soma em relação ao índice especificado:

Distribuição Conjunta de X e Y

	y_1	y_2	y_3	...	y_n	$p_{i.}$
x_1	p_{11}	p_{12}	p_{13}	...	p_{1n}	$p_{1.}$
x_2	p_{21}	p_{22}	p_{23}	...	p_{2n}	$p_{2.}$
x_3	p_{31}	p_{32}	p_{33}	...	p_{3n}	$p_{3.}$
...	...	...	...		...	...
x_m	p_{m1}	p_{m2}	p_{m3}	...	p_{mn}	$p_{m.}$
$p_{.j}$	$p_{.1}$	$p_{.2}$	$p_{.3}$	...	$p_{.n}$	1

Donde as várias notações usadas:

$$p_{i.} = P(x_i) = P(X = x_i) = \sum_j P(X = x_i; Y = y_j) = \sum_j p_{ij} \quad i = 1,2,3,\ldots,m$$

$$p_{.j} = P(y_j) = P(Y = y_j) = \sum P(X = x_i; Y = y_j) = \sum_i p_{ij} \quad j = 1,2,3,\ldots,n$$

Média de X e média de Y

$$\mu_X = E(X) = x_1 p_{1.} + x_2 p_{2.} + x_3 p_{3.} + \cdots + x_m p_{m.}$$

$$\mu_Y = E(Y) = y_1 p_{.1} + y_2 p_{.2} + y_3 p_{.3} + \cdots + y_n p_{.n}$$

que são **parâmetros de posição** das distribuições de X e Y respectivamente.

Variância de X e variância de Y

$$\sigma_X^2 = Var(X) = E(X - \mu_X)^2 = (x_1 - \mu_X)^2 p_{1.} + (x_2 - \mu_X)^2 p_{2.} + \cdots + (x_m - \mu_X)^2 p_{m.}$$

$$\sigma_Y^2 = Var(Y) = E(Y - \mu_Y)^2 = (y_1 - \mu_Y)^2 p_{.1} + (y_2 - \mu_Y)^2 p_{.2} + \cdots + (y_n - \mu_Y)^2 p_{.n}$$

são **parâmetros de dispersão** das distribuições de X e Y respectivamente.

Covariância de X e Y e Correlação linear de X e Y

$$\sigma_{XY}^2 = Cov(X,Y) = E(X-\mu_X)(Y-\mu_Y)$$

$$= (x_1-\mu_X)(y_1-\mu_Y)p_{11} + (x_1-\mu_X)(y_2-\mu_Y)p_{2n} + \cdots + (x_1-\mu_X)(y_n-\mu_Y)p_{1n} +$$

$$(x_2-\mu_X)(y_1-\mu_Y)p_{21} + (x_2-\mu_X)(y_2-\mu_Y)p_{22} + \cdots + (x_2-\mu_X)(y_n-\mu_Y)p_{2n} +$$

$$\ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots$$

$$(x_m-\mu_X)(y_1-\mu_Y)p_{m1} + (x_m-\mu_X)(y_2-\mu_Y)p_{2n} + \cdots + (x_m-\mu_X)(y_n-\mu_Y)p_{mn}$$

é a covariância de X e Y e

$$\rho_{XY} = \frac{\sigma_{XY}^2}{\sigma_X \sigma_Y}$$

é a **correlação linear** de X e Y, que são **parâmetros de associação** das distribuições de X e Y. Não é difícil demonstrar que $-1 \le \rho_{XY} \le 1$.

Exemplo de distribuições com mesma dispersão e médias diferentes

x	$P_1(X)$	$P_2(X)$
1	0,1	0
2	0,2	0
3	0,4	0
4	0,2	0,1
5	0,1	0,2
6	0	0,4
7	0	0,2
8	0	0,1
	1	1

$$\mu_1 = 1\cdot 0{,}1 + 2\cdot 0{,}2 + 3\cdot 0{,}4 + 4\cdot 0{,}2 + 5\cdot 0{,}1 = 3$$

$$\mu_2 = 4\cdot 0{,}1 + 5\cdot 0{,}2 + 6\cdot 0{,}4 + 7\cdot 0{,}2 + 8\cdot 0{,}1 = 6$$

$$\sigma_X^2 = (1-3)^2\cdot 0{,}1 + (2-3)^2\cdot 0{,}2 + (3-3)^2\cdot 0{,}4 + (4-3)^2\cdot 0{,}2 + (5-3)^2\cdot 0{,}1 = 1{,}2$$

$$\sigma_Y^2 = (4-6)^2\cdot 0{,}1 + (5-6)^2\cdot 0{,}2 + (6-6)^2\cdot 0{,}4 + (7-6)^2\cdot 0{,}2 + (8-6)^2\cdot 0{,}1 = 1{,}2$$

Percebe-se que P_2 está apenas deslocada à direita em relação a P_1. A distância média ao quadrado dos valores de X e Y relativos às suas médias permanecem iguais.

Exemplo de distribuições com mesma média e dispersões diferentes

x	$P_3(X)$	$P_4(X)$
2	0,1	0
3	0,2	0,2
4	0,4	0,6
5	0,2	0,2
6	0,1	0
	1	1

$$\mu_3 = 2 \cdot 0{,}1 + 3 \cdot 0{,}2 + 4 \cdot 0{,}4 + 5 \cdot 0{,}2 + 6 \cdot 0{,}1 = 4$$

$$\mu_4 = 2 \cdot 0{,}0 + 3 \cdot 0{,}2 + 4 \cdot 0{,}6 + 5 \cdot 0{,}2 + 6 \cdot 0{,}0 = 4$$

$$\sigma_3^2 = (2-4)^2 \cdot 0{,}1 + (3-4)^2 \cdot 0{,}2 + (4-4)^2 \cdot 0{,}4 + (5-4)^2 \cdot 0{,}2 + (6-4)^2 \cdot 0{,}1 = 1{,}2$$

$$\sigma_4^2 = (2-4)^2 \cdot 0{,}0 + (3-4)^2 \cdot 0{,}2 + (4-4)^2 \cdot 0{,}6 + (5-4)^2 \cdot 0{,}2 + (6-4)^2 \cdot 0{,}0 = 0{,}4$$

Nesses casos a distribuição P_3 e P_4 localizam-se na mesma posição (médias iguais) mas P_4 concentra mais os valores em torno de sua média quando comparada a P_3.

Exemplo de cálculo de covariância e correlação linear

$P_1(x|y)$, $P_1(x)$ e $P_1(y)$

x \ y	**1**	**2**	**3**	**4**	**5**	**P(x)**
1	0	0	0,0	0,5	1	**0,2**
3	0	0	0,5	0,5	0	**0,3**
5	0	0,5	0,5	0	0	**0,3**
7	1	0,5	0,0	0	0	**0,2**
P(y)	**0,1**	**0,2**	**0,4**	**0,2**	**0,1**	**1**

Nota: Como a tabela acima contém as distribuições condicionais de X dado Y=y a coluna de P(x) é obtida pela fórmula da probabilidade total, ou seja é a média das distribuições condicionais ponderada pela marginal de Y.

As médias e variâncias de X e Y serão:

$$\mu_X = 1\cdot 0{,}2 + 3\cdot 0{,}3 + 5\cdot 0{,}3 + 7\cdot 0{,}2 = 4$$

$$\mu_Y = 1\cdot 0{,}1 + 2\cdot 0{,}2 + 3\cdot 0{,}4 + 4\cdot 0{,}2 + 5\cdot 0{,}1 = 3$$

$$\sigma_X^2 = (1-4)^2\cdot 0{,}2 + (3-4)^2\cdot 0{,}3 + (5-4)^2\cdot 0{,}3 + (7-4)^2\cdot 0{,}2 = 4{,}2$$

$$\sigma_Y^2 = (1-3)^2\cdot 0{,}1 + (2-3)^2\cdot 0{,}2 + (3-3)^2\cdot 0{,}4 + (4-3)^2\cdot 0{,}2 + (5-3)^2\cdot 0{,}1 = 1{,}2$$

A distribuição conjunta de X e Y fica :

$P_1(x,y)$, $P_1(x)$ e $P_1(y)$

x \ y	**1**	**2**	**3**	**4**	**5**	**P(x)**
1	0	0	0	0,1	0,1	**0,2**
3	0	0	0,2	0,1	0	**0,3**
5	0	0,1	0,2	0	0	**0,3**
7	0,1	0,1	0	0	0	**0,2**
P(y)	**0,1**	**0,2**	**0,4**	**0,2**	**0,1**	**1**

$$\begin{aligned}\sigma_{XY}^2 = {} & (1-4)(4-3)\cdot 0{,}1 + (1-4)(5-3)\cdot 0{,}1 + \\ & +(3-4)(3-3)\cdot 0{,}2 + (3-4)(4-3)\cdot 0{,}1 + \\ & +(5-4)(2-3)\cdot 0{,}1 + (5-4)(3-3)\cdot 0{,}2 + \\ & +(7-4)(1-3)\cdot 0{,}1 + (7-4)(2-3)\cdot 0{,}1 = -2\end{aligned}$$

donde

$$\rho_{XY} = \frac{\sigma_{XY}^2}{\sigma_X\sigma_Y} = \frac{-2}{\sqrt{4{,}2\cdot 1{,}2}} = \frac{-2}{2{,}245} = -0{,}89$$

16. Propriedades da Média e da Variância de uma distribuição

Propriedades da Média

Teorema: $E(X+Y) = E(X) + E(Y)$ para todo par de variáveis aleatórias X e Y:

Prova: Pela definição de esperança matemática temos:

$$E(X) = x_1 p_{1.} + x_2 p_{2.} + \cdots + x_m p_{m.} \ , \ E(Y) = y_1 p_{.1} + y_2 p_{.2} + \cdots + y_n p_{.n}$$

e

$$\begin{aligned} E(X+Y) = \ & (x_1+y_1)p_{11} + (x_1+y_2)p_{12} + \cdots + (x_1+y_n)p_{1n} + \\ & (x_2+y_1)p_{21} + (x_2+y_2)p_{22} + \cdots + (x_2+y_n)p_{2n} + \\ & + \ \ldots \ \ldots \ \ldots \ \ldots \ \ldots \ \ldots \ \ldots \ \ldots \ + \\ & (x_m+y_1)p_{m1} + (x_m+y_2)p_{m2} + \cdots + (x_m+y_n)p_{mn} \end{aligned}$$

Distribuindo a soma e usando a definição de $p_{i.}$ temos:

$$\begin{aligned} E(X+Y) = \ & x_1 p_{1.} + y_1 p_{11} + y_2 p_{12} + \cdots + y_n p_{1n} + \\ & x_2 p_{2.} + y_1 p_{21} + y_2 p_{22} + \cdots + y_n p_{2n} + \\ & + \ \ldots \ \ldots \ \ldots \ \ldots \ \ldots \ \ldots \ + \\ & x_m p_{m.} + y_1 p_{m1} + y_2 p_{m2} + \cdots + y_n p_{mn} \end{aligned}$$

Fatorando y_j e usando a definição de $p_{.j}$ temos:

$$E(X+Y) = x_1 p_{1.} + x_2 p_{2.} + \cdots + x_m p_{m.} + y_1 p_{.1} + y_2 p_{.2} + \cdots + y_n p_{.n} = E(X) + E(Y)$$

cqd.

Corolário: $E(X_1 + X_2 + \cdots + X_m) = E(X_1) + E(X_2) + \cdots + E(X_m)$

Prova: Imediata por indução completa.

Teorema: Se X_1 e X_2 são independentes então $E(X \cdot Y) = E(X) \cdot E(Y)$.

Comentário: Como por hipótese X e Y são independentes poderemos usar a fatoração da distribuição conjunta: $p_{ij} = p_{i.}\, p_{.j}$ para todo par (i , j), isto é, a distribuição conjunta é o produto das marginais. É importante deixar claro que esse teorema tem aplicação muito menos abrangente do que o da soma, que não impôs a condição de independência. Em amostras independentes esse teorema será muito usado.

Prova:

$$\begin{aligned} E(X \cdot Y) = \ & (x_1 \cdot y_1)p_{11} + (x_1 \cdot y_2)p_{12} + \cdots + (x_1 \cdot y_n)p_{1n} + \\ & (x_2 \cdot y_1)p_{21} + (x_2 \cdot y_2)p_{22} + \cdots + (x_2 \cdot y_n)p_{2n} + \\ & + \ \ldots \ \ldots \ \ldots \ \ldots \ \ldots \ \ldots \ \ldots \ + \\ & (x_m \cdot y_1)p_{m1} + (x_m \cdot y_2)p_{m2} + \cdots + (x_m \cdot y_n)p_{mn} \end{aligned}$$

$$\begin{aligned} E(X \cdot Y) = \ & (x_1 \cdot y_1)p_{1.}p_{.1} + (x_1 \cdot y_2)p_{1.}p_{.2} + \cdots + (x_1 \cdot y_n)p_{1.}p_{.n} + \\ & (x_2 \cdot y_1)p_{2.}p_{.1} + (x_2 \cdot y_2)p_{2.}p_{.2} + \cdots + (x_2 \cdot y_n)p_{2.}p_{.n} + \\ & + \ \ldots \ \ldots \ \ldots \ \ldots \ \ldots \ \ldots \ \ldots \ \ldots \ \ldots \ \ldots \ + \\ & (x_m \cdot y_1)p_{m.}p_{.1} + (x_m \cdot y_2)p_{m.}p_{.2} + \cdots + (x_m \cdot y_n)p_{m.}p_{.n} \end{aligned}$$

$$\begin{aligned} E(X \cdot Y) = \; & x_1 \cdot p_{1.}(y_1 \cdot p_{.1} + y_2 \cdot p_{.2} + \cdots + y_n \cdot p_{.n}) + \\ & x_2 \cdot p_{2.}(y_1 \cdot p_{.1} + y_2 \cdot p_{.2} + \cdots + y_n \cdot p_{.n}) + \\ & + \; \ldots \; \ldots \; \ldots \; \ldots \; \ldots \; \ldots \; \ldots \; + \\ & x_m \cdot p_{m.}(y_1 \cdot p_{.1} + y_2 \cdot p_{.2} + \cdots + y_n \cdot p_{.n}) \end{aligned}$$

$$E(X \cdot Y) = (x_1 \cdot p_{1.} + x_2 \cdot p_{2.} + \cdots + x_m \cdot p_{m.})(y_1 \cdot p_{.1} + y_2 \cdot p_{.2} + \cdots + y_n \cdot p_{.n})$$

$$E(X \cdot Y) = E(X) \cdot E(Y)$$ **cqd.**

Corolário 1: Se X e Y são independentes então $E[R(X) \cdot S(Y)] = E[R(X)] \cdot E[S(Y)]$

Prova: Fazendo $U = R(X)$ e $V = S(X)$, recaindo no teorema anterior. **cqd**

Corolário 2: Se $X_1, X_2, \cdots, X_m$ são independentes então

$$E(X_1 \cdot X_2 \cdot \;\cdots\; \cdot X_m) = E(X_1) \cdot E(X_2) \cdot \;\cdots\; \cdot E(X_m)$$

Prova: Imediata por indução completa.

Comentário: Mesmo quando X e Y são independentes não são todas as funções de X e Y que permitem obter a esperança matemática por fatoração. Isso só acontece quando a função também fatorar, como no caso dos corolários.

Teorema: Se **C** é uma constante **c** com probabilidade 1 então $E(C) = c$.

Prova: $E(C) = c \cdot P(C = c) = c \cdot 1 = c$ **cqd**.

Teorema: $E(c \cdot X) = c \cdot E(X)$

Prova: $E(c \cdot X) = c \cdot x_1 p_{1.} + c \cdot x_2 p_{2.} + \cdots + c \cdot x_m p_{m.} = c \cdot \sum_{i=1}^{m} x_i p_{i.} = c \cdot E(X)$ **cqd**.

Reunindo os teoremas temos

$E(C) = c$	$E(c \cdot X) = c \cdot E(X)$	$E(X + Y) = E(X) + E(Y)$

e, sob independência de X e Y

$E(X \cdot Y) = E(X) \cdot E(Y)$

Exemplo: Sejam $X_1, X_2, \cdots, X_m$ idêntica e independente distribuídas como **X** com média **μ**, então $E(\overline{X}) = \mu$, onde $\overline{X} = \frac{1}{m}(X_1 + X_2 + \cdots + X_m)$. Isto é, a esperança matemática da média amostral é igual à média populacional quando a amostra é completamente aleatória.

Prova: Usando os teoremas acima demonstrados temos:

$$E(\overline{X}) = E\left[\frac{1}{m}(X_1 + X_2 + \cdots + X_m)\right] = \frac{1}{m}E(X_1 + X_2 + \cdots + X_m) = \frac{1}{m}(EX_1 + EX_2 + \cdots + EX_m)$$

$$= \frac{1}{m}(\mu + \mu + \cdots + \mu) = \frac{m \cdot \mu}{m} = \mu$$

Exemplo: Se **K** tem distribuição binomial com parâmetros **n** e **p** então $E(K) = np$.

Prova: Sejam $B_1, B_2, \ldots, B_n$ variáveis indicadoras da ocorrência ou não do evento **E** nas sucessivas provas de Bernoulli tratadas na distribuição binomial:

$$\left\{\begin{matrix} P(B_1 = 1) = p \\ P(B_1 = 0) = 1 - p = q \end{matrix}\right\}, \left\{\begin{matrix} P(B_2 = 1) = p \\ P(B_2 = 0) = 1 - p = q \end{matrix}\right\}, \ldots, \left\{\begin{matrix} P(B_n = 1) = p \\ P(B_n = 0) = 1 - p = q \end{matrix}\right\}$$

donde, $E(B_j) = 1 \cdot p + 0 \cdot q = p \; j = 1, 2, \cdots, n$

dessa forma o número de ocorrências do evento **N** será dado por $K = B_1 + B_2 + \ldots + B_n$ donde, aplicando o teorema da esperança da soma teremos:

$$E(K) = E(B_1 + B_2 + \cdots + B_n) = E(B_1) + E(B_2) + \cdots + E(B_n) = n \cdot p$$ **cqd**.

Propriedades da Variância

Teorema: Se **C** é constante com probabilidade **1** então **Var(C)=0**

Prova: $Var(C) = E[C - E(C)]^2 = (c - c)^2 \cdot 1 = 0$ **cqd**.

Teorema: Se **C** é constante então $Var(X + C) = Var(X)$

Prova: Como $E(X + C) = E(X) + E(C) = E(X) + C$ então

$$Var(X+C)=E(X+C-E(X+C))^2=E(X+C-E(X)-C)^2=E(X-E(X))^2=Var(X)$$

cqd.

Teorema: $Var(X)=E(X^2)-\mu_X^2$

Prova:

$$Var(X)=E(X-\mu_X)^2=E(X^2-2\mu_X X+\mu_X^2)=E(X^2)-2\mu_X E(X)+E(\mu_X^2)$$

$$=E(X^2)-2\mu_X\mu_X+\mu_X^2=E(X^2)-\mu_X^2$$ **cqd**.

Corolário: Do último resultado resulta imediatamente:

$$E(X^2)=Var(X)+\mu_X^2=\sigma_X^2+\mu_X^2=Var(X)+[E(X)]^2$$

Teorema: $Var(c\cdot X)=c^2\cdot Var(X)$

Prova: Como $E(c\cdot X)=c\cdot E(X)$ e $E(c\cdot X)^2=E(c^2X^2)=c^2E(X^2)$, aplicando o último teorema para a variável aleatória $c\cdot X$ temos:

$$Var(c\cdot X)=E(c\cdot X)^2-[E(c\cdot X)]^2=c^2\cdot E(X^2)-c^2[E(X)]^2$$

$$=c^2\left\{E(X^2)-[E(X)]^2\right\}=c^2\cdot Var(X)$$

cqd

Teorema: Se $X\ e\ Y$ são independentes então:

$$Var(X+Y)=Var(X)+Var(Y)$$

Prova: Pela hipótese de independência, $E(X\cdot Y)=E(X)\cdot E(Y)=\mu_X\mu_Y$. Então,

$$Var(X+Y)=E(X+Y)^2-\mu_{X+Y}^2=E\left(X^2+2X\cdot Y+Y^2\right)-(\mu_X+\mu_Y)^2$$

$$=EX^2+2E(X\cdot Y)+EY^2-\mu_X^2-2\mu_X\mu_Y-\mu_Y^2$$

$$=EX^2+2\mu_X\mu_Y+EY^2-\mu_X^2-2\mu_X\mu_Y-\mu_Y^2$$

$$=EX^2-\mu_X^2+EY^2-\mu_Y^2=Var(X)+Var(Y)$$

cqd.

Corolário: Se $X_1,X_2,\cdots,X_m$ são independentes então

$$Var(X_1+X_2+\cdots+X_m)=Var(X_1)+Var(X_2)+\cdots+Var(X_m)$$

Prova: (Exercício de indução completa)

Reunindo os teoremas sobre variância temos

$\boxed{Var(C)=0}$ $\boxed{Var(c\cdot X)=c^2\cdot Var(X)}$ $\boxed{Var(X+c)=Var(X)}$ $\boxed{Var(X)=E(X^2)-\mu_X^2}$

e, sob independência de $X_1,X_2,\cdots X_m$

$$\boxed{Var(X_1 + X_2 + \cdots + X_m) = Var(X_1) + Var(X_2) + \cdots + Var(X_m)}$$

Teorema: Seja m $X_1, X_2, \cdots, X_m$ cópias idênticas e independentes de X, uma variável aleatória com variância σ_X^2. Seja $\overline{X}$ a média amostral, então $\sigma_{\overline{X}}^2 = \frac{\sigma_X^2}{m}$.

Prova:

$$\sigma_{\overline{X}}^2 = Var\left(\frac{X_1 + X_2 + \cdots + X_m}{m}\right) = Var\left(\frac{X_1}{m}\right) + Var\left(\frac{X_2}{m}\right) + \cdots + Var\left(\frac{X_m}{m}\right)$$

$$= \frac{1}{m^2}Var(X_1) + \frac{1}{m^2}Var(X_2) + \cdots + \frac{1}{m^2}Var(X_m)$$

$$= \frac{1}{m^2}\sigma_X^2 + \frac{1}{m^2}\sigma_X^2 + \cdots + \frac{1}{m^2}\sigma_X^2 = \frac{m\sigma_X^2}{m^2} = \frac{\sigma_X^2}{m}$$

cqd.

Exemplo: Se **K** tem distribuição Binomial com parâmetros **n** e **p** então $Var(K) = npq$

Prova: $Var(K) = Var(B_1 + B_2 + \cdots B_n) = Var(B_1) + Var(B_2) + \cdots + Var(B_n) = nVar(B_1)$

mas $Var(B_1) = (1-p)^2 p + (0-p)^2 q = q^2 p + p^2 q = pq(p+q) = pq$, donde, $Var(K) = npq$

cqd.

Reunindo os resultados dos dois exemplos temos que se **K** tem **distribuição binomial** com parâmetros **n** e **p** então

$$\boxed{E(K) = n \cdot p} \quad \text{e} \quad \boxed{Var(K) = npq}$$

Distribuição de Poisson (Freqüência de evento raro)

Como já vimos a variável binomial **K** representa a freqüência de ocorrência de um evento qualquer **E** numa seqüência de Bernoulli de tamanho **n**. Quando a probabilidade **p** é pequena e o tamanho da amostra **n** grande (>30) então **K** é a freqüência de evento raro e sua distribuição se aproxima da distribuição limite $P(K=k) = \frac{\lambda^k}{k!}e^{-\lambda}, k = 0,1,2,\cdots$ chamada **Distribuição de Poisson**. Lembrando que $e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \cdots$, não é muito difícil demonstrar que

$$\lim_{\substack{n\to\infty \\ p\to 0 \\ \lambda = np}} C_n^k p^k q^{n-k} = \frac{\lambda^k}{k!}e^{-\lambda}, k = 0,1,2,\cdots$$

Exemplo: Se **K** tem distribuição de Poisson com parâmetro λ então $\mu_K = \lambda$ e $\sigma_K^2 = \lambda$.

Prova: Na demonstração anterior temos que $E(K) = np = \lambda$ fixo na operação de limite. Ademais, como $p \to 0$, $q = 1 - p \to 1$, então $Var(K) = npq \to np = \lambda$ também.

Exemplo: Numa certa região do Semi-árido nordestino o número médio λ de dias com mais de 5mm, de chuva é de **1**. Obtenha a distribuição do número **K** de dias com mais de 5mm. de chuva.

Solução: Como se trata de um evento raro podemos usar a distribuição de Poisson, resultando:

$$P(K = k) = \frac{\lambda^k}{k!} e^{-\lambda} \; k = 0,1,2,\cdots \text{ onde } \lambda = 1$$

k	P(k)
0	0,367879
1	0,367879
2	0,183940
3	0,061313
4	0,015328
5	0,003066
6	0,000511
...	...

17. Estimação da Média e da Variância

Teorema: Sejam $X_1, X_2, \cdots, X_n$ uma amostra completamente aleatória de X com média μ e variância σ^2 .

Sejam as estatísticas $\overline{X}$ e S^2 definidas pelas expressões:

$$\overline{X} = \frac{X_1 + X_2 + \cdots + X_n}{n} \qquad\qquad S^2 = \frac{1}{n-1}\left(\sum_i^n X_i^2 - n\overline{X}^2\right)$$

então $E(\overline{X}) = \mu$ e $E(S^2) = \sigma^2$.

Comentário: Demonstrado esse teorema estaremos diante de dois **estimadores imparciais** de μ e σ^2, respectivamente. Salientamos que se trata de um resultado bem geral aplicado a qualquer variável aleatória.

Prova: Como $E(\overline{X}) = \mu$ já ficou demonstrado anteriormente, resta a prova de que $E(S^2) = \sigma^2$. Lembremos ainda que já provamos que $\mu_{\overline{X}} = \mu$ e $\sigma^2_{\overline{X}} = \dfrac{\sigma^2}{n}$. Então temos:

$$E(S^2) = E\left[\frac{1}{n-1}\left(\sum_i^n X_i^2 - n\overline{X}^2\right)\right] =$$

$$\frac{1}{n-1}\left[\left(\sum_i^n EX_i^2 - nE\overline{X}^2\right)\right] =$$

mas $EX_i^2 = \sigma^2 + \mu^2$ e $E\overline{X}^2 = \sigma^2_{\overline{X}} + \mu^2_{\overline{X}} = \dfrac{\sigma^2}{n} + \mu^2$ resultando

$$E(S^2) = \frac{1}{n-1}\left[\left(\sum_i^n (\sigma^2 + \mu^2) - n(\frac{\sigma^2}{n} + \mu^2)\right)\right] =$$

$$E(S^2) = \frac{1}{n-1}\left[n(\sigma^2 + \mu^2) - \sigma^2 - n\mu^2\right] = \frac{1}{n-1}\left[(n-1)\sigma^2\right] = \sigma^2$$

cqd

As estatísticas $\overline{X}$ e S^2 são chamadas respectivamente **média amostral** e **variância amostral** em contraposição a μ e σ^2 que são chamadas respectivamente **média populacional** e **variância populacional**.

Amostra	$\overline{x}_5$	$\overline{x}_{10}$	s_5^2	s_{10}^2
1	2,8	3,0	1,20	0,66
2	2,8	2,6	0,70	1,15
3	2,4	2,6	1,30	1,15
4	3,0	2,7	2,00	2,01
5	3,8	3,7	0,12	1,02
6	2,8	3,2	2,70	0,40
7	3,4	3,4	0,80	0,71
8	3,6	3,4	1,20	2,60
[Min ; Max]	[2,4 ; 3,8]	[2,6 ; 3,7]	[0,12 ; 2,70]	[0,40 ; 2,60]

A tabela no fim da página anterior dispõe 8 amostras obtidas por alunos do Curso de Estatística de uma variável aleatória **X** com a seguinte distribuição e parâmetros:

$$P(X=k)=\frac{k}{10}k=1,2,3,4$$

$$\mu=E(X)=1\cdot\frac{1}{10}+2\cdot\frac{2}{10}+3\cdot\frac{3}{10}+4\cdot\frac{4}{10}=\frac{30}{10}=3$$

$$\sigma^2=Var(X)=(1-3)^2\cdot\frac{1}{10}+(2-3)^2\cdot\frac{2}{10}+(3-3)^2\cdot\frac{3}{10}+(4-3)^2\cdot\frac{4}{10}=\frac{10}{10}=1$$

Observem que as médias amostrais ambas oscilam em torno do parâmetro $\mu_X=3$ e as variâncias amostrais observadas oscilam em torno de $\sigma_X^2=1$. Em ambos os casos as amostras de tamanho 10 tem dispersão menor comparadas as de tamanho 5 em relação aos parâmetros que estimam. Essa é uma simulação de cunho didático visando mostrar empiricamente a imparcialidade dos estimadores e suas precisões em função do tamanho da amostra.

18. Distribuições Contínuas

Considere X uma variável aleatória aplicada em Ω assumindo valores em **R** e que

$$P(X\in[a,b])=\int_a^b f(x)dx \text{ onde } f(x)\geq 0 \text{ e } \int_{-\infty}^{+\infty} f(x)dx=1$$

Então X é definida como **variável aleatória contínua**, $f(x)$ **a função de densidade** e $f(x)dx$ a **diferencial de probabilidade** de X. Nesse contexto teremos as definições:

$$\mu=E(X)=\int_{-\infty}^{\infty} xf(x)dx \qquad E(X^2)=\int_{-\infty}^{\infty} x^2 f(x)dx \qquad Var(X)=\int_{-\infty}^{\infty}(x-\mu)^2 f(x)dx$$

Estendendo para um par de variáveis aleatórias teremos que se (X,Y) for aplicada a Ω assumindo valores em $\mathbf{R^2}$ e

$P[(X,Y)\in[a,b]\times[c,d]]=\int_a^b\int_c^d f(x,y)dydx$ onde $f(x,y)\geq 0$ e $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dydx=1$ então (X,Y) tem uma

distribuição
conjunta de probabilidade contínua com **função de densidade** $f(x,y)$ e **X** e **Y** são independentes se

$$f_{XY}(x,y)=f_X(x)f_Y(y) \quad \forall x,y$$

Nesse contexto teremos as definições:

$$E(X\cdot Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x\cdot y\cdot f(x,y)dxdy \qquad Cov(X\cdot Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x-\mu_X)(y-\mu_Y)\cdot f(x,y)dxdy$$

Observação Importante: Todas as propriedades de média e variância estendem-se para o caso contínuo, pois a integral é o limite de uma soma, requerendo apenas de cuidados de rigor matemático para serem provadas.

Exemplo: Um cabo de aço de 2 metros preso nas extremidades é esticado até a ruptura. Se o cabo for homogêneo, qual a probabilidade desse ponto de ruptura dar-se entre 0,5m e 1,5m de uma das extremidades? Calcular $E(X)$ e $Var(X)$.

Solução: É realístico considerar a função de densidade

$$\begin{cases} f(x) = \dfrac{1}{2} & x \in [0,2] \\ f(x) = 0 & x \notin [0,2] \end{cases}$$

donde,

$$P(0,5 \leq X \leq 1,5) = \int_{0,5}^{1,5} \frac{1}{2}\, dx = \frac{1}{2}[x]_{0,5}^{1,5} = \frac{1}{2}(1,5-0,5) = \frac{1}{2}$$

$$\mu = E(X) = \int_{0,5}^{1,5} x \cdot \frac{1}{2} dx = \frac{1}{2}\left[\frac{x^2}{2}\right]_0^2 = \frac{1}{4}(4-0) = 1$$

$$E(X^2) = \int_0^2 x^2 \cdot \frac{1}{2} dx = \frac{1}{2}\left[\frac{x^3}{3}\right]_0^2 = \frac{1}{6}(8-0) = \frac{4}{3}$$

$$Var(X) = E(X^2) - \mu^2 = \frac{4}{3} - 1 = \frac{1}{3}$$

Uma variável como do último exemplo é chamada de **distribuição uniforme no intervalo [0 , 2]**. Podemos generalizar a **distribuição uniforme no intervalo [a , b]** com a função de densidade:

$$\begin{cases} f(x) = \dfrac{1}{b-a} & x \in [a,b] \\ f(x) = 0 & x \notin [a,b] \end{cases}$$

Exercício: Considere X com distribuição uniforme em [a , b]. Calcule E(X) e Var(X).

Resposta: $E(X) = \dfrac{a+b}{2}$ e $Var(X) = \dfrac{(a-b)^2}{12}$

Exercício: Seja

$$f(x) = \begin{cases} 6x(1-x) & x \in [0;1] \\ 0 & x \notin [0;1] \end{cases}$$

a função de densidade de **X**. Mostre que essa função atende às condições de ser uma função de densidade (**não negativa** e integral em **R** igual a **1**); Calcule a média μ e a variância σ^2.

Respostas: A função dada é uma parábola truncada com raízes 0 e 1 com valores positivos entre as raízes;

$$\int_0^1 6x(1-x)dx = 1\ ;\ \mu = E(X) = \int_0^1 x6x(1-x)dx = \frac{1}{2}\ ;\ E(X^2) = \int_0^1 x^2 6x(1-x)dx = \frac{3}{10}\quad \sigma^2 = E(X^2) - \mu^2 = \frac{1}{20}\,.$$

19. O Teorema do Limite Central e a Distribuição Normal

Teorema: Sejam $X_1, X_2, \cdots, X_n, \cdots$ uma amostra completamente aleatória de X de tamanho crescente tal que $\mu = E(X)$ e $\sigma^2 = Var(X)$. Então

$$\sqrt{n}\left(\frac{\overline{X}-\mu}{\sigma}\right) = \frac{\overline{X}-\mu}{\sqrt{\dfrac{\sigma^2}{n}}} = \frac{\overline{X}-\mu}{\sqrt{\sigma^2_{\overline{X}}}} = \frac{\overline{X}-\mu}{\sigma_{\overline{X}}}$$

tem uma distribuição assintótica de probabilidade $(n \to \infty)$ contínua como **Z** de função de densidade dada por

$$p(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}\ z \in R\,.$$

Observação: Essa distribuição, sem qualquer dúvida é a mais importante das distribuições e largamente empregada em todas as ciências, matemáticas ou não e foi obtida de forma independente por Gauss e Laplace e denominada **Distribuição Normal Padrão**, completamente tabelada. Para usarmos no Curso, usaremos uma tabela simplificada com apenas uma casa decimal.

Tabela Simplificada da Distribuição Normal Padrão			
z	$T(z) = P(0 \le Z \le z) = \int_0^z p(z)dz$	z	$T(z) = P(0 \le Z \le z) = \int_0^z p(z)dz$
0,0	0,0000	2,0	0,4772
0,1	0,0398	2,1	0,4821
0,2	0,0793	2,2	0,4861
0,3	0,1179	2,3	0,4893
0,4	0,1554	2,4	0,4918
0,5	0,1915	2,5	0,4938
0,6	0,2258	2,6	0,4953
0,7	0,2580	2,7	0,4965
0,8	0,2881	2,8	0,4974
0,9	0,3159	2,9	0,4981
1,0	0,3413	3,0	0,4987
1,1	0,3643	3,1	0,4990
1,2	0,3849	3,2	0,4993
1,3	0,4032	3,3	0,4995
1,4	0,4192	3,4	0,4997
1,5	0,4332	3,5	0,4998
1,6	0,4452	3,6	0,4998
1,7	0,4554	3,7	0,4999
1,8	0,4641	3,8	0,4999
1,9	0,4713	3,9	0,5000

Exemplo: Seja **K** com distribuição binomial com parâmetros **n** e **p**. Como

$$K = B_1 + B_2 + \cdots + B_n$$

onde $B_1, B_2, \cdots, B_n$ são idêntica e independente distribuídas como **B** sendo $\mu_B = p$ e $\sigma_B^2 = pq$ então

$\sqrt{n}\,\dfrac{\frac{K}{n} - p}{\sqrt{p(1-p)}} = \dfrac{K - np}{\sqrt{np(1-p)}}$ tem distribuição assintótica $(n \to \infty)$ normal padrão, como dita o Teorema do Limite Central.

Exemplo: Sejam $U_1, U_2, \cdots, U_n$ uma amostra completamente aleatória (significando variáveis idêntica e independentemente distribuídas) de uma variável **U** uniforme no intervalo [0 ; 1] . Isto é, a função de densidade de **U** será

$$p(u) = \begin{cases} 1 & u \in [0;1] \\ 0 & u \notin [0;1] \end{cases}$$

Pelo resultado de exercício dado, $\mu_U = E(U) = \frac{1+0}{2} = \frac{1}{2}$ e $\sigma_U^2 = Var(U) = \frac{(1-0)^2}{12} = \frac{1}{12}$. A média amostral $\overline{U} = \frac{U_1 + U_2 + \cdots + U_n}{n}$ tem média $\mu_{\overline{U}} = \frac{1}{2}$ e variância $\sigma_{\overline{U}}^2 = Var(\overline{U}) = \frac{1}{12n}$. Então $\frac{\overline{U} - \frac{1}{2}}{\sqrt{\frac{1}{12n}}}$ tem distribuição assintótica $(n \to \infty)$ normal padrão. Multiplicando e dividindo por **n** a expressão resulta em $\frac{\sum_{i=1}^{n} U_i - \frac{n}{2}}{\sqrt{\frac{n}{12}}}$ com distribuição assintótica $(n \to \infty)$ normal padrão.

20. Média e Variância da Distribuição Normal Padrão

$\mu_Z = E(Z) = \int_{-\infty}^{\infty} z \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$ Como trata-se de uma função ímpar [f(-z) = - f(z)] essa integral é nula, donde E(Z) = 0. $E(Z^2) = \int_{-\infty}^{\infty} z^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = 2\int_{0}^{\infty} z^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$ pelo fato da função integranda ser par [f(-z) = f(z)] .

Fazendo $t = \frac{z^2}{2}$, $z = \sqrt{2} \cdot t^{\frac{1}{2}}$, $dt = z \cdot dz$ e $dz = \frac{1}{\sqrt{2}} t^{-\frac{1}{2}} dt$ resulta

$E(Z^2) = 2\int_{0}^{\infty} 2t \frac{1}{\sqrt{2\pi}} e^{-t} \frac{1}{\sqrt{2}} t^{-\frac{1}{2}} dt = \frac{2}{\sqrt{\pi}} \int_{0}^{\infty} t^{\frac{1}{2}} \cdot e^{-t} dt$. Mas do cálculo integral sabemos que, pelas propriedades da função gama, $\int_{0}^{\infty} t^{\frac{1}{2}} \cdot e^{-t} dt = \Gamma(\frac{1}{2} + 1) = \frac{1}{2}\Gamma(\frac{1}{2}) = \frac{\sqrt{\pi}}{2}$. Portanto,

$E(Z^2) = \frac{2}{\sqrt{\pi}} \frac{\sqrt{\pi}}{2} = 1$. Portanto, $\sigma_Z^2 = Var(Z) = E(Z^2) - \mu_Z^2 = 1 - 0^2 = 1$.

Resumindo diremos que se **Z** tem distribuição normal padrão então $\mu_z = 0$ e $\sigma_z^2 = 1$. É comum a notação **Z** ∩ **N [0 ; 1]** . Nesse contexto o símbolo ∩ é lido como "**distribuído como**".

Distribuição Normal Geral Y ∩ N [μ ; σ²]

Fazendo $Y = \sigma \cdot Z + \mu$ resulta $\mu_Y = \mu$ e $\sigma_Y^2 = \sigma^2$. Substituindo $z = \frac{y-\mu}{\sigma}$ e $dz = \frac{1}{\sigma}dy$ na densidade da normal padrão o elemento de probabilidade de **Y** fica

$$p(y)dy = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y-\mu)^2}{2\sigma^2}} dy y \in R \ ,$$

que se define como a função de densidade da **distribuição normal geral**, com média μ e variância σ^2. Portanto os parâmetros de uma distribuição normal são a média e a variância dela.

Quando temos Y $\cap$ N [μ ; σ^2], efetuando a transformação $Z = \frac{Y-\mu}{\sigma}$ **,** recaímos na variável normal padrão, totalmente tabelada. Veremos agora como usar a tabela de T(z) para calcular certas probabilidades envolvendo variáveis normais.

Exemplo: Sabe-se que a produção **Y** de banana numa certa região tem média μ = 10t/ha e um desvio padrão σ de 2 t/ha . Calcular $P(Y > 12)$; $P(Y < 8)$; $P(7 < Y < 13)$; $P(6 \le Y < 8)$ e $P(13 < Y < 15)$.

Respostas: O problema fornece os valores $\mu = 10$ e $\sigma = 2$. Usando a transformação-z temos, considerando a simetria da função de densidade normal padrão:

$$P(Y > 12) = P(\frac{Y-\mu}{\sigma} > \frac{12-10}{2}) = P(Z > 1) = 0{,}5 - T(1) = 0{,}5 - 0{,}3413 = 0{,}1587$$

$$P(Y < 8) = P(\frac{Y-\mu}{\sigma} < \frac{8-10}{2}) = P(Z < -1) = P(Z > 1) = 0{,}5 - T(1) = 0{,}5 - 0{,}3413 = 0{,}1587$$

$$P(7 < Y < 13) = P(\frac{7-10}{2} < \frac{Y-\mu}{\sigma} < \frac{13-10}{2}) =$$

$$P(-1{,}5 < Z < 1{,}5) = P(-1{,}5 < Z < 0) + P(0 < Z < 1{,}5) = 2 \cdot T(1{,}5) = 2 \cdot 0{,}4332 = 0{,}8664$$

$$P(6 < Y < 8) = P(\frac{6-10}{2} < \frac{Y-\mu}{\sigma} < \frac{8-10}{2}) =$$

$$P(-2 < Z < -1) = P(1 < Z < 2) = T(2) - T(1) = 0{,}4772 - 0{,}3413 = 0{,}1359$$

$$P(13 < Y < 15) = P(\frac{13-10}{2} < \frac{Y-\mu}{\sigma} < \frac{15-10}{2}) =$$

$$P(1{,}5 < Z < 2{,}5) = T(2{,}5) - T(1{,}5) = 0{,}4938 - 0{,}4332 = 0{,}0606$$

Exercício: No exemplo anterior considere que, após um manejo melhor, essa cultura passou a produzir 0,5 t/ha a mais. Considerando que o desvio padrão não se alterou, obtenha as probabilidades dos mesmos eventos pedidos.

Exercício: Calcule $P(Y > 3)$ para Y$\cap$ N[0 ,1]; Y$\cap$ N[0,5 ; 1]; Y$\cap$ N[1 ; 1] e Y$\cap$ N[2 ; 1] .

21. Distribuição da soma de variáveis normais independentes

Sejam $X \cap N[\mu;\sigma^2]$ e $Y \cap N[\mu;\sigma^2]$ e independentes. O elemento de probabilidade conjunta será

$$dp(x,y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \, \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y-\mu)^2}{2\sigma^2}} dy \, .$$

Se efetuarmos a transformação $u = x + yv = x - y$ teremos $x = \frac{u+v}{2} y = \frac{u-v}{2}$. Então

$$\det\begin{bmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{bmatrix} = -1/2$$

é o Jacobiano e o elemento de probabilidade de **u** e **v** fica

$$dp(u,v) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(\frac{u+v}{2}-\mu)^2}{2\sigma^2}} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(\frac{u-v}{2}-\mu)^2}{2\sigma^2}} \frac{1}{2} dudv$$

Somando os expoentes e simplificando, obtemos:

$$dp(u,v) = \frac{1}{\sqrt{2}\sigma\sqrt{\pi}} e^{-\frac{(u-2\mu)^2}{2(2\sigma^2)}} du \frac{1}{\sqrt{2}\sigma\sqrt{\pi}} e^{-\frac{v^2}{2(2\sigma^2)}} dv$$

$$dp(u,v) = dp(u) \cdot dp(v)$$

Portanto **U = X + Y** tem distribuição normal **N [2μ ; 2σ^2]** e **U = X - Y** tem distribuição normal **N [0 ; 2σ^2]** . Ademais **U** e **V** são independentes.

22. Função Geradora de Momentos

A função $M_x(t) = E(e^{tX})$,quando existe, caracteriza uma distribuição. Ela é muito importante para distribuições contínuas e, para o caso da normal é especialmente interessante:

$$M_Z(t) = E(e^{tZ}) = \int_{-\infty}^{\infty} e^{tz} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}+tz} dz$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2-2tz}{2}} dz = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2-2tz+t^2-t^2}{2}} dz$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(z-t)^2 - t^2}{2}} dz = e^{\frac{t^2}{2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(z-t)^2}{2}} dz$$

Como a função integranda é uma densidade normal **N [t ; 1]** temos $M_Z(t) = e^{\frac{t^2}{2}}$. Agora se $Y = \sigma \cdot Z + \mu$ então

$$M_Y(t) = E(e^{tY}) = E\left[e^{t(\sigma \cdot Z + \mu)}\right] = e^{t\mu} E\left[e^{(\sigma \cdot t)Z}\right]$$

Portanto a função geradora de momentos da normal geral fica

$$M_Y(t) = e^{t\mu} M_Z(\sigma \cdot t) = e^{t\mu} e^{\frac{\sigma^2 t^2}{2}} = e^{\mu \cdot t + \frac{\sigma^2 t^2}{2}}$$

Exemplo: Uma aplicação importante é a generalização da distribuição da soma de duas normais independentes quaisquer. Senão vejamos: Se $X_1 \cap N\,[\mu_1;\sigma_1^2]$ e $X_2 \cap N\,[\mu_2;\sigma_2^2]$ então, devido à independência,

$$M_{X_1 + X_2}(t) = Ee^{t(X_1 + X_2)} = Ee^{tX_1} e^{tX_2} = Ee^{tX_1} Ee^{tX_2} = M_{X_1}(t) M_{X_2}(t)$$

$$M_{X_1 + X_2}(t) = e^{\mu_1 \cdot t + \frac{\sigma_1^2 t^2}{2}} e^{\mu_2 \cdot t + \frac{\sigma_2^2 t^2}{2}} = e^{(\mu_1 + \mu_2) \cdot t + \frac{(\sigma_1^2 + \sigma_2^2) t^2}{2}}$$

que é a função geradora de momentos de uma normal $N[\mu_1 + \mu_2; \sigma_1^2 + \sigma_2^2]$, isto é, a soma de duas normais independentes também é normal.

Exercício: Mostre, usando a função geradora de momentos, que se

$$X_1 \cap N\,[\mu_1;\sigma_1^2],\ X_2 \cap N\,[\mu_2;\sigma_2^2],\quad \ldots \quad,\ X_n \cap N\,[\mu_n;\sigma_n^2]$$

então

$$c_1 X_1 + c_2 X_2 + \cdots + c_n X_n \cap N\left[c_1\mu_1 + c_2\mu_2 + \cdots + c_n\mu_n; c_1^2\sigma_1^2 + c_2^2\sigma_2^2 + \cdots + c_n^2\sigma_n^2\right].$$

23. Tamanho de uma amostra

Uma das primeiras questões que se coloca quando se aborda estimação de parâmetros de uma distribuição é qual o número **n** de unidades experimentais devem-se para que a precisão de um estimador seja adequada. Vamos dar uma resposta simples e eficiente para essa questão.

Sejam $X_1, X_2, \cdots, X_n, \cdots$ uma amostra de tamanho crescente de **X** de média **μ** e variância **σ²**. Pelo teorema do limite central $\dfrac{\overline{X}-\mu}{\sqrt{\dfrac{\sigma^2}{n}}} \underset{n\to\infty}{\cap} N[0,1]$. Convém afirmar aqui que se a distribuição de **X** for simétrica, um valor apenas maior que 10 já resulta numa distribuição da média amostral muito próxima da normal. Caso as distribuições das variáveis já sejam normal então $\dfrac{\overline{X}-\mu}{\sqrt{\dfrac{\sigma^2}{n}}} \cap N[0,1]$ de forma exata. Em ambos os casos, pela tabela de normal padrão, podemos escrever

$$P(-2 < \frac{\overline{X}-\mu}{\sqrt{\frac{\sigma^2}{n}}} < 2) \cong 0{,}95 \text{ ou } P\{|\overline{X}-\mu| < 2\sqrt{\frac{\sigma^2}{n}}\} \cong 0{,}95$$

Chamando $d = 2\sqrt{\dfrac{\sigma^2}{n}}$ a **margem de erro**, obtemos $d^2 = 4\dfrac{\sigma^2}{n} \Rightarrow n = 4\dfrac{\sigma^2}{d^2}$. Observe que o tamanho é diretamente proporcional à variância dos dados e inversamente proporcional ao quadrado da margem de erro. Como em geral não sabemos o valor de **σ²**, usamos a estimativa imparcial **s²**, resultando uma estimativa $\mathbf{n_0}$ de **n**.

$$n_0 = 4\frac{s^2}{d^2}$$

Para obter-se **s²** devemos ter uma amostra piloto ou inicial e $\mathbf{n_0}$ será o tamanho da amostra inicial. Numa segunda etapa da pesquisa um novo valor de $\mathbf{n_0}$ pode ser recalculado.

24. Intervalo de Confiança para a média de X

Para **n** grande (>30) podemos sem grande erro substituir **σ²** por **S²**, obtendo a expressão:

$$P\left\{\left|\overline{X}-\mu\right| < 2\sqrt{\frac{S^2}{n}}\right\} \cong 0{,}95 \Leftrightarrow P\left\{\left|\mu-\overline{X}\right| < 2\sqrt{\frac{S^2}{n}}\right\} \cong 0{,}95$$

Dessa expressão decorre que $\left[\overline{X}-2\sqrt{\frac{S^2}{n}};\overline{X}+2\sqrt{\frac{S^2}{n}}\right]$ tem aproximadamente **0,95** de probabilidade de incluir

μ . Evidentemente, numa amostra específica, o intervalo

$$IC_{0,95}(\mu)=\left[\bar{x}-2\sqrt{\frac{s^2}{n}};\bar{x}+2\sqrt{\frac{s^2}{n}}\right]$$

observado não é mais aleatório e dizemos que é um **intervalo de confiança ao nível de 0,95** da média μ . O termo confiança substitui probabilidade para atender ao rigor matemático.

Exemplo: A freqüência relativa $f=\overline{B}=\frac{K}{n}$ de uma seqüência de provas de Bernoulli tem média p e

variância $\frac{pq}{n}$. Nesse caso f é claramente $\overline{X}$, agora , para **n** grande:

$$S^2=\frac{1}{n-1}\left[B_1^2+B_2^2+\cdots+B_n^2-n\overline{B}^2\right]=\frac{1}{n-1}(nf-nf^2)=\frac{nf(1-f)}{n-1}$$

donde,

$$IC_{0,95}(p)=\left[f-2\sqrt{\frac{f(1-f)}{n-1}};f+2\sqrt{\frac{f(1-f)}{n-1}}\right]$$

é um intervalo de confiança de **p** ao nível de 0,95.

25. Intervalo de Confiança para Diferenças (Uma amostra e grande)

Sejam $(X_1,Y_1),(X_2,Y_2),\cdots,(X_n,Y_n)$ pares de dados de **n** indivíduos com a distribuição de (X,Y). Portanto as variáveis X e

Y não são necessariamente independentes mas a amostra é completamente aleatória. Sejam os dados

$D_1=X_1-Y_1, D_2=X_2-Y_2,\cdots,D_n=X_n-Y_n$ réplicas tais que $\mu_D=E(D)=E(X-Y)=\mu_X-\mu_Y$ e

$\sigma_D^2=Var(D)=Var(X-Y)$. Pelo Teorema do Limite Central, fazendo $\overline{D}=\overline{X}-\overline{Y}$ e

$$S_D^2 = \frac{1}{n-1}\left[D_1^2 + D_2^2 + \cdots + D_n^2 - n\overline{D}^2\right],$$

teremos:

$$IC_{0,95}(\mu_X - \mu_Y) = \left[\overline{x} - \overline{y} - 2\sqrt{\frac{s_D^2}{n}}; \overline{x} - \overline{y} + 2\sqrt{\frac{s_D^2}{n}}\right]$$

um intervalo de confiança para μ_x-μ_y ao nível de 0,95. Esse resultado pode ser aplicado a todo ensaio do tipo "antes" e "depois" de um tratamento aplicado a **n** indivíduos.

26. Distribuição de t de Student: Essa seção trata de amostras de variáveis normais pequenas **(n≤30)**. W. Gosset (Student) trabalhando de forma empírica no controle de qualidade de uma cervejaria obteve por simulação a tabela da variável

$$T = \sqrt{n}\,\frac{\overline{X} - \mu}{S}$$

de uma amostra completamente aleatória de $X \cap N(\mu;\sigma^2)$. A tabela simplificada para efeitos didáticos fornece os valores críticos $\mathbf{t_{0,95}}$ tais que $P(-t_{0,95} < T < t_{0,95}) = 0{,}95$.

Limites Críticos Bilaterais a 95% da Distribuição T de Student			
n-1	$\mathbf{t_{0,95}}$	**n-1**	$\mathbf{t_{0,95}}$
1	12,706	**16**	2,120
2	4,313	**17**	2,110
3	3,182	**18**	2,101
4	2,776	**19**	2,093
5	2,571	**20**	2,086
6	2,447	**21**	2,080
7	2,365	**22**	2,074
8	2,306	**23**	2,069
9	2,262	**24**	2,064
10	2,228	**25**	2,060
11	2,201	**26**	2,056
12	2,179	**27**	2,052
13	2,160	**28**	2,048
14	2,145	**29**	2,045
15	2,131	**30**	2,042

Exemplo: A tabela seguinte exibe na sua primeira coluna 15 observações de peso de 100 grãos de sementes de feijão. Vamos calcular três intervalos de confiança com tamanhos crescentes para mostrar o aumento da precisão em função do tamanho da amostra.

Dados de Feijão e Cálculo de I.C. para Três Tamanhos de Amostra			
x_i	$\sum x_i$	x_i^2	$\sum x_i^2$
14		196	
18		324	
15		225	
21		441	
19	87	361	1547
16		256	
17		289	
22		484	
18		324	
14	174	196	3096
17		289	
17		289	
13		169	
20		400	
19	260	361	4604

Cálculo dos Intervalos de Confiança:

$$IC_{0,95}(\mu) = \frac{87}{5} \pm 2{,}776\sqrt{\frac{\frac{1}{4}\left[1547 - 5\left(\frac{87}{5}\right)^2\right]}{5}} = 17{,}4 \pm 3{,}58 \text{ com as primeiras 5 observações}$$

$$IC_{0,95}(\mu) = \frac{174}{10} \pm 2{,}262\sqrt{\frac{\frac{1}{9}\left[3096 - 10\left(\frac{174}{10}\right)^2\right]}{10}} = 17{,}4 \pm 1{,}97 \text{ com as primeiras 10 observações}$$

$$IC_{0,95}(\mu) = \frac{260}{15} \pm 2{,}145\sqrt{\frac{\frac{1}{14}\left[4604 - 15\left(\frac{260}{15}\right)^2\right]}{15}} = 17{,}3 \pm 1{,}46 \text{ com todas as 15 observações}$$

Exemplo de Cálculo de Tamanho da Amostra: Considere a amostra com 10 observações. A variância amostral dá $s^2 = 7{,}6$. Se quisermos D = 1,46 (precisão do IC com 15 observações) então

$$n_0 = \frac{4 \cdot s^2}{d^2} = \frac{4 \cdot 7{,}6}{1{,}46^2} = 14{,}3$$

indicando um tamanho de amostra de 14 observações. Na verdade com 15 obtemos essa precisão. Mas não deixa de ser uma boa aproximação.

26. Distribuição de Qui-quadrado (χ^2) de Pearson

Seja $Z \cap N(0;1)$ e sejam $Z_1, Z_2, \cdots, Z_n$ uma amostra completamente aleatória de **Z**. Então $X_r^2 = Z_1^2 + Z_2^2 + \cdots + Z_r^2$ tem distribuição Qui-quadrado com **r** gráus de liberdade. O elemento de probabilidade é

$$dp_{X^2}(x) = \frac{1}{\Gamma(r/2) \cdot 2^{r/2}} x^{\frac{r}{2}-1} e^{-\frac{x}{2}} dx, x > 0$$

Limites Críticos Unilaterais a 95% do Qui-quadrado de Pearson em Função do Número de Gráus de Liberdade			
r	$\mathbf{x^2_{0,95}}$	**r**	$\mathbf{x^2_{0,95}}$
1	3,84	**16**	26,3
2	5,99	**17**	27,6
3	7,81	**18**	28,9
4	9,49	**19**	30,1
5	11,1	**20**	31,4
6	12,6	**21**	32,7
7	14,1	**22**	33,9
8	15,5	**23**	35,2
9	16,9	**24**	36,4
10	18,3	**25**	37,7
11	19,7	**26**	38,9
12	21,0	**27**	40,1
13	22,4	**28**	41,3
14	23,7	**29**	42,6
15	25,0	**30**	43,8

Teste de Significância Estatística: Aplicação do Qui-quadrado: Uma Categoria Duas Classes (1 gráu de liberdade)

Como já vimos anteriormente,

$$\frac{f-p}{\sqrt{\frac{p(1-p)}{n}}}$$

é aproximadamente normal padrão quando **n** é grande. Então o seu quadrado será aproximadamente Qui-quadrado com 1 gráu de liberdade. Com um pouco de álgebra podemos nos convencer que

$$\frac{n(f-p)^2}{p(1-p)}=\frac{(nf-np)^2}{np(1-p)}=\frac{(nf-np)^2}{np}+\frac{[n(1-f)-n(1-p)]^2}{n(1-p)}\cap \mathrm{X}_1^2$$

Fazendo $F_1 = nf$ a freqüência bruta observada da ocorrência do evento **E**, $F_2 = n(1-f)$ a freqüência bruta observada da não ocorrência de **E**, $T_1 = np$ a freqüência bruta esperada teórica da ocorrência de **E** e $T_2 = n(1-p)$ a freqüência bruta esperada teórica da não ocorrência de **E**, teremos, quando **n** é grande

$$\frac{(F_1-T_1)^2}{T_1}+\frac{(F_2-T_2)^2}{T_2}\cap \mathrm{X}_1^2$$

Define-se $\left\{\mathrm{X}_1^2 > x_{1-\alpha}^2\right\}$como **região de rejeição** e $P\left\{\mathrm{X}_1^2 > x_{1-\alpha}^2\right\}$ como **nível de significância (α)** da hipótese : H_0: $E(F_1) = T_1$, $E(F_2) = T_2$. Nos exemplos usaremos $1-\alpha = 0{,}95$.

Exemplo: Uma planta **Aa** e outra **aa** para um locus são cruzadas . Pela teoria espera-se que os descendentes desse cruzamento sejam 50% do tipo **Aa** e 50% do tipo **aa**. Realizado um experimento observou-se, de 120 descendentes 53 do tipo **Aa** e 67 do tipo **aa**. Verifique pelo Qui-quadrado se a hipótese **H_0: $E(F_1) = E(F_2) = 60$** pode ser sustentada por esses dados.

Solução: Pela expressão acabada de ser deduzida temos:

$$\frac{(53-60)^2}{60}+\frac{(67-60)^2}{60}=\frac{49+49}{60}=\frac{98}{60}=1{,}63$$

Usando $X_1^2 > 3{,}84$ como região de rejeição para a hipótese, concluímos que não devemos rejeitar a teoria usada; as freqüências observadas "aderem" às esperadas o suficiente. Esse tipo de teste é também conhecido como **Teste de Aderência**.

Uma Categoria e Mais de Uma Classe (r = n_c-1 gráus de liberdade)

Se tivermos $\mathbf{n_c}$ classes de uma categoria então a expressão anterior generaliza-se para

$$\frac{(F_1-T_1)^2}{T_1}+\frac{(F_2-T_2)^2}{T_2}+\cdots+\frac{(F_{n_c}-T_{n_c})^2}{T_{n_c}} \cap X^2_{n_c-1}$$

Exemplo: A produção de milho é dada aproximadamente pelo modelo matemático $S = 20 + X + Y + R$ com X, Y, R independentes onde as distribuições de X, Y e R são dadas por:

x	**P(x)**	**y**	**P(y)**	**r**	**P(r)**
-2	1/5	-3	1/3	-1	1/2
0	3/5	0	1/3	1	1/2
2	1/5	3	1/3		
Total	1		1		1

Combinando os 3x3x2 possíveis pontos amostrais teremos:

Espaço Amostral , Valores da Produção e Probabilidades		
ϖ (x, y, r)	**s = 20 + x + y + r**	**P(ϖ)**
(-2, -3, -1)	14	1/30
(-2, -3,+1)	16	1/30
(-2, 0, -1)	17	1/30
(-2, 0,+1)	19	1/30
(-2,+3, -1)	20	1/30
(-2,+3,+1)	22	1/30
(0, -3, -1)	16	3/30
(0, -3,+1)	18	3/30
(0, 0, -1)	19	3/30
(0, 0, +1)	21	3/30
(0, +3, -1)	22	3/30
(0, +3,+1)	24	3/30
(+2, -3, -1)	18	1/30
(+2, -3,+1)	20	1/30
(+2, 0, -1)	21	1/30
(+2, 0,+1)	23	1/30
(+2,+3, -1)	24	1/30
(+2,+3,+1)	26	1/30
Total		**1**

Uma amostra de 120 produções foi tomada e a distribuição de freqüências das produções foram as da tabela abaixo:

Classe	**s**	**F_i**
1	14	4
2	16	17
3	17	6
4	18	14
5	19	15
6	20	10
7	21	17
8	22	14
9	23	5
10	24	13
11	26	5
	Total	**120**

Será que essas freqüências brutas observadas "aderem" às esperadas pela teoria ? Pelos dados do exemplo podemos obter a distribuição de **S,** a produção de milho. Multiplicando por 120 obtemos as freqüências brutas esperadas teóricas e teremos todos os dados para aplicar o Qui-quadrado. O número **r** de gráus de liberdade do Qui-quadrado será $\mathbf{n_c}$**-1**, o número de classes da categoria produção menos um, ou seja **11 - 1 = 10**.

Distribuição de S e Cálculo do Qui-quadrado				
s	**P(s)**	**F_i**	**T_i**	$\frac{(F_i - T_i)^2}{E_i}$
14	1/30	**4**	**4**	0,0000
16	4/30	**17**	**16**	0,0625
17	1/30	**6**	**4**	1,0000
18	4/30	**14**	**16**	0,2500
19	4/30	**15**	**16**	0,0625
20	2/30	**10**	**8**	0,5000
21	4/30	**17**	**16**	0,0625
22	4/30	**14**	**16**	0,2500
23	1/30	**5**	**4**	0,2500
24	4/30	**13**	**16**	0,5625
26	1/30	**5**	**4**	0,2500
Total	**1**	**120**	**120**	**3,2500**

O total da última coluna **supostamente** nos dá o valor observado de um Qui-quadrado teórico com 10 gráus de liberdade. Da tabela temos que o valor crítico para rejeitar ou não a hipótese é **18,3**. Portanto, como **3,25 < 18,3** concluímos que as freqüências observadas aderem às esperadas o suficiente para **não rejeitarmos** o modelo matemático proposto no exemplo.

Duas Categorias e Mais de Uma Classe [r = (n_l-1)(n_c-1) gráus de liberdade]

Exemplo: Numa pesquisa sobre o estado geral do rebanho caprino em Salvador foi obtida a tabela abaixo exibindo as freqüências brutas de 490 animais classificados em duas categorias: Tipo de Manejo e Idade. Quer-se testar se essas freqüências observadas diferem significativamente de freqüências esperadas, sob a hipótese de independência das categorias. A hipótese de independência das categorias impõe a igualdade das probabilidades condicionais às respectivas marginais em ambos os sentidos e, entre parênteses, estão as freqüências esperadas sob essa hipótese. Por exemplo:

$$\frac{(T_{11})}{113} = \frac{47}{490} \Rightarrow (T_{11}) = 10{,}8 \qquad \frac{(T_{12})}{113} = \frac{85}{490} \Rightarrow (T_{12}) = 19{,}6 \qquad \frac{(T_{13})}{113} = \frac{358}{490} \Rightarrow (T_{13}) = 82{,}6$$

$$\frac{(T_{21})}{56} = \frac{47}{490} \Rightarrow (T_{21}) = 5{,}4 \qquad \frac{(T_{22})}{56} = \frac{85}{490} \Rightarrow (T_{22}) = 9{,}7 \qquad \frac{(T_{23})}{56} = \frac{358}{490} \Rightarrow (T_{23}) = 40{,}9$$

$$\frac{(T_{31})}{93} = \frac{47}{490} \Rightarrow (T_{31}) = 8{,}9 \qquad \frac{(T_{32})}{93} = \frac{85}{490} \Rightarrow (T_{32}) = 16{,}1 \qquad \frac{(T_{33})}{93} = \frac{358}{490} \Rightarrow (T_{33}) = 67{,}9$$

... ...

etc.

Freqüências observadas (F_{ij}) e esperadas (T_{ij}) de 490 caprinos classificados por Manejo e Idade
Parcelas da Soma do Cálculo do Qui-quadrado com (5-1)(3-1) = 8 g.l.
Hipótese sendo testada: Independência entre Manejo e Idade

Idade \ Manejo	1	2	3	Total
1	**10** (10,8) 0,0592	**15** (19,6) 1,0796	**88** (82,6) 0,3530	**113**
2	**5** (5,4) 0,0296	**28** (9,7) 34,5247	**23** (40,9) 7,8340	**56**
3	**21** (8,9) 16,4506	**14** (16,1) 0,2739	**58** (67,9) 1,4434	**93**
4	**7** (10,6) 1,2226	**18** (19,1) 0,0634	**85** (80,4) 0,2632	**110**
5	**4** (11,3) 4,7159	**10** (20,5) 5,3780	**104** (86,2) 3,6756	**118**
Total	**47**	**85**	**358**	**490**

Da tabela de valores críticos do Qui-quadrado temos $P(X_8^2 > 15{,}5) = 0{,}05$. Mas, como podemos conferir pelos cálculos intermediários exibidos na tabela acima,

$$\sum_{i=1}^{5}\sum_{j=1}^{3}\frac{(F_{ij}-T_{ij})^2}{T_{ij}} = 77{,}4$$

Portanto o Qui-quadrado observado cai na região de rejeição, pois 77,4 > 15.5. Em outras palavras, sob independência das categorias, o evento $X_8^2 > 15{,}5$, raro (P=0,05), ocorreu. Preferimos, pois, descartar a hipótese de independência. Para confirmar essa conclusão o leitor deve percorrer a tabela acima e verificar que, para algumas caselas, as freqüências observadas e esperadas diferem mais do que o razoável.

Grupo de Problemas II

1. Em levantamento efetuado na região de Salvador 195 caprinos foram classificados em duas categorias: Sexo (linhas) e Índice Clínico (Colunas). Considerando os dados da tabela de contingência abaixo teste a hipótese H_0 de Independência entre as categorias.
2.

Sexo \ **Índice Clínico**	**I**	**II**	**III**	**Total**
Fêmeas	53	74	36	**163**
Machos	6	16	10	**32**
Total	**59**	**90**	**46**	**195**

Resposta: $X_6^2 = 2{,}75$ (não significativo, hipótese não rejeitada)

3. Dados da produção (X e Y) de soja em duas regiões foram obtidos. Sejam :

$$X \cap N(\mu_1;\sigma_1^2) \quad \text{e} \quad Y \cap N(\mu_2;\sigma_2^2)$$

onde as variâncias são desconhecidas. Usando o T de Student obtenha os intervalos de confiança respectivos e conclua se as médias são estatisticamente diferentes, usando o seguinte critério: Se os intervalos tiverem pontos comuns então as médias não diferem, de outra forma conclua que elas diferem.

x	20	17	23	19	21	20	20	18	22	20
y	25	21	29	20	30	22	28	24	26	25

Resposta:

$$IC_{0,95}(\mu_1) = 20 \pm 2{,}262\sqrt{\frac{3{,}11}{10}} = 20 \pm 1{,}26 = [18{,}74;21{,}26]$$

$$IC_{0,95}(\mu_2) = 25 \pm 2{,}262\sqrt{\frac{11{,}33}{10}} = 25 \pm 2{,}41 = [22{,}59\,;27{,}41]$$

Como os intervalos de confiança não possuem pontos comuns conclui-se que, estatisticamente, as médias diferem entre si ao nível de confiança de 95%.

4. A concentração de creatina (T) no sangue, por pesquisas já efetuadas, tem uma distribuição dada pela função de densidade:

$$f(t) = \begin{cases} \dfrac{3t^2}{1000} & t \in [0;10] \\ 0 & t \notin [0;10] \end{cases}$$

   a) Prove que f(t) é uma função de densidade.
   b) Calcule E(T) e Var(T).
   c) Calcule $F(x) = P(T \leq x) x \in R$ e faça o gráfico de F(x).
   d) Calcule M_d tal que $F(M_d) = 0{,}5$, a mediana de T.

Respostas:

a) A integral de f em **R** é 1 b) E(T) = 7,5 Var(T) = 3,75

c) $$F(x) = \begin{cases} 0 & x < 0 \\ \dfrac{x^3}{1000} & x \in [0;10] \\ 1 & x > 10 \end{cases}$$; d) $M_d = \sqrt[3]{500} = 7{,}9$

5. Cinco candidatos (A, B, C, D e E) disputam uma eleição. Um instituto de pesquisa de opinião sorteia ao acaso na população 701 eleitores, obtendo os seguintes resultados:

CANDIDATO	PORCENTAGEM
A	30
B	28
C	25
D	10
E	7
TOTAL	100

Obter o Intervalo de Confiança a 95% para cada candidato e discutir a classificação estatística obtida.

Resposta:

$$IC_{0,95}(p_A) = 0,30 \pm 0,035 = [0,265\,;\,0,335]$$

$$IC_{0,95}(p_B) = 0,28 \pm 0,034 = [0,246\,;\,0,314]$$

$$IC_{0,95}(p_C) = 0,25 \pm 0,033 = [0,217\,;\,0,283]$$

$$IC_{0,95}(p_D) = 0,10 \pm 0,023 = [0,077;\,0,123]$$

$$IC_{0,95}(p_E) = 0,07 \pm 0,019 = [0,051;\,0,089]$$

Classificação Estatística:

E (7%) D (10%) C (25%) B (28%) A (30%)

Conclusão: Os candidatos unidos por uma linha não diferem estatisticamente e aqueles não ligados diferem estatisticamente a 95% de confiança.

6. Calcule $P(Y > 3)$ para $Y \cap N[0\,,1]$; $Y \cap N[0,5\,;\,1]$; $Y \cap N[1\,;\,1]$ e $Y \cap N[2\,;\,1]$

Respostas:

Se $Y \cap N(0\,;\,1)$ então $P(Y > 3) = 0,0013$

$$\text{Se } Y \cap N(0{,}5\ ;\ 1) \quad \text{então} \quad P(Y > 3) = 0{,}0062$$

$$\text{Se } Y \cap N(1\ ;\ 1) \quad \text{então} \quad P(Y > 3) = 0{,}0228$$

$$\text{Se } Y \cap N(2\ ;\ 1) \quad \text{então} \quad P(Y > 3) = 0{,}1587$$

7. Seja $Y_{ij} \cap N(i, j)$ para $i = 1{,}2$ e $j = 1{,}2{,}3$ independentes. Obtenha as distribuições de

a) $Y_{1.}$ e $Y_{2.}$ b) $Y_{.1}$, $Y_{.2}$ e $Y_{.3}$ c) $\overline{Y}_{1.}$ e $\overline{Y}_{2.}$ d) $\overline{Y}_{.1}$, $\overline{Y}_{.2}$ e $\overline{Y}_{.3}$

onde

$$\overline{Y}_{i.} = \frac{Y_{i.}}{3} \quad \text{e} \quad \overline{Y}_{.j} = \frac{Y_{.j}}{2}$$

Solução: Na forma de tabela temos:

i \ j	1	2	3
1	$Y_{11} \cap N(1\ ;\ 1)$	$Y_{12} \cap N(1\ ;\ 2)$	$Y_{13} \cap N(1\ ;\ 3)$
2	$Y_{21} \cap N(2\ ;\ 1)$	$Y_{22} \cap N(2\ ;\ 2)$	$Y_{23} \cap N(2\ ;\ 3)$

$$E(Y_{1.}) = E(Y_{11} + Y_{12} + Y_{13}) = E(Y_{11}) + E(Y_{12}) + E(Y_{13}) = 1 + 1 + 1 = 3$$

$$Var(Y_{1.}) = Var(Y_{11} + Y_{12} + Y_{13}) = Var(Y_{11}) + Var(Y_{12}) + Var(Y_{13}) = 1 + 2 + 3 = 6$$

$$E(Y_{2.}) = E(Y_{21} + Y_{22} + Y_{23}) = E(Y_{21}) + E(Y_{22}) + E(Y_{23}) = 2 + 2 + 2 = 6$$

$$Var(Y_{2.}) = Var(Y_{21} + Y_{22} + Y_{23}) = Var(Y_{21}) + Var(Y_{22}) + Var(Y_{23}) = 1 + 2 + 3 = 6$$

Portanto,

$$Y_{1.} \cap N(3\ ;\ 6) \quad \text{e} \quad Y_{2.} \cap N(6\ ;\ 6)$$

$$E(Y_{.1}) = E(Y_{11} + Y_{21}) = E(Y_{11}) + E(Y_{21}) = 1 + 2 = 3$$

$$Var(Y_{.1}) = Var(Y_{11} + Y_{21}) = Var(Y_{11}) + Var(Y_{21}) = 1 + 1 = 2$$

$$E(Y_{.2}) = E(Y_{12} + Y_{22}) = E(Y_{12}) + E(Y_{22}) = 1 + 2 = 3$$

$$Var(Y_{.2}) = Var(Y_{12} + Y_{22}) = Var(Y_{12}) + Var(Y_{22}) = 2 + 2 = 4$$

$$E(Y_{.3}) = E(Y_{13} + Y_{23}) = E(Y_{13}) + E(Y_{23}) = 1 + 2 = 3$$

$$Var(Y_{.3}) = Var(Y_{13} + Y_{23}) = Var(Y_{13}) + Var(Y_{23}) = 3 + 3 = 6$$

$$Y_{.1} \cap N(3;2)\ ,\ Y_{.2} \cap N(3;4)\ \text{ e }\ Y_{.3} \cap N(3;6)$$

$$E(\overline{Y}_{1.}) = E\left(\frac{1}{3}Y_{1.}\right) = \frac{1}{3}E(Y_{1.}) = \frac{1}{3}(3) = 1 \qquad Var(\overline{Y}_{1.}) = Var\left(\frac{1}{3}Y_{1.}\right) = \frac{1}{9}Var(Y_{1.}) = \frac{6}{9} = \frac{2}{3}$$

$$E(\overline{Y}_{2.}) = E\left(\frac{1}{3}Y_{2.}\right) = \frac{1}{3}E(Y_{2.}) = \frac{6}{3} = 2 \qquad Var(\overline{Y}_{2.}) = Var\left(\frac{1}{3}Y_{2.}\right) = \frac{1}{9}Var(Y_{2.}) = \frac{6}{9} = \frac{2}{3}$$

$$E(\overline{Y}_{.1}) = E\left(\frac{1}{2}Y_{.1}\right) = \frac{1}{2}E(Y_{.1}) = \frac{3}{2} \qquad Var(\overline{Y}_{.1}) = Var\left(\frac{1}{2}Y_{.1}\right) = \frac{1}{4}Var(Y_{.1}) = \frac{2}{4} = \frac{1}{2}$$

$$E(\overline{Y}_{.2}) = E\left(\frac{1}{2}Y_{.2}\right) = \frac{1}{2}E(Y_{.2}) = \frac{3}{2} \qquad Var(\overline{Y}_{.2}) = Var(\frac{1}{2}Y_{.2}) = \frac{1}{4}Var(Y_{.2}) = \frac{4}{4} = 1$$

$$E(\overline{Y}_{.3}) = E\left(\frac{1}{2}Y_{.3}\right) = \frac{1}{2}E(Y_{.3}) = \frac{3}{2} \qquad Var(\overline{Y}_{.3}) = Var(\frac{1}{2}Y_{.3}) = \frac{1}{4}Var(Y_{.3}) = \frac{6}{4} = \frac{3}{2}$$

Portanto,

$$\overline{Y}_{.1} \cap N(\frac{3}{2};\frac{1}{2})\ ,\ \overline{Y}_{.2} \cap N(\frac{3}{2};1)\ \text{ e }\ \overline{Y}_{.3} \cap N(\frac{3}{2};\frac{3}{2})$$

8. O tempo de espera Y para atendimento de 430 clientes de um caixa de banco foi classificado em classes de intervalos e são apresentados abaixo. Nesse caso, como os dados originais não existem, as fórmulas para a obtenção de $\overline{y}$ e s^2 são:

$$\overline{y} = \frac{F_1 \cdot m_1 + F_2 \cdot m_2 + \cdots + F_I \cdot m_I}{n}$$

$$s^2 = \frac{1}{n-1}\left(F_1 \cdot m_1^2 + F_2 \cdot m_2^2 + \cdots + F_I \cdot m_I^2 - n \cdot \overline{y}^2\right)$$

onde $\mathbf{F_i}$ e $\mathbf{m_i}$ são a freqüência observada e o ponto médio do intervalo **i**, **n** o número de observações e **I** o número de intervalos usado.

a) Calcule $\overline{y}$ e s^2 para esses dados.

b) Obtenha as probabilidades de **Y** pertencer a cada intervalo, caso $Y \cap N(\overline{y};s^2)$.

c) Calcule a coluna de $\mathbf{T_i}$ de freqüências teóricas esperadas.

d) Calcule o Qui-quadrado, que nesse caso terá **I - 2** gráus de liberdade (dois parâmetros estimados). Teste a 0,95 de significância se os dados são normais.

Intervalo	m_i	F_i
$(-\infty;13]$	12	18
$[13;15]$	14	48
$[15;17]$	16	93
$[17;19]$	18	104
$[19;21]$	20	100
$[21;23]$	22	40
$[23;+\infty)$	24	27
Total		430

Respostas:

a)

$$\bar{y} = \frac{18 \cdot 12 + 48 \cdot 14 + 93 \cdot 16 + 104 \cdot 18 + 100 \cdot 20 + 40 \cdot 22 + 27 \cdot 24}{430} = \frac{7776}{430} = 18{,}1$$

$$s^2 = \frac{1}{429}\left(18 \cdot 12^2 + 48 \cdot 14^2 + 93 \cdot 16^2 + 104 \cdot 18^2 + 100 \cdot 20^2 + 40 \cdot 22^2 + 27 \cdot 24^2 - 430 \cdot 18{,}1^2\right) = \frac{144416 - 140872{,}3}{429} = 8{,}26$$

b) e c) Como $s = \sqrt{8{,}26} \cong 2{,}9$ e usando a transformação $z_i = \frac{m_i - \bar{y}}{s} i = 1, 2, 3, 4, 5, 6 \; e \; 7$, temos:

Intervalo (Y)	Intervalo (Z)	P_i	m_i	F_i	T_i (430 . P_i)
$(-\infty\ ;13]$	$(-\infty\ ;-1{,}8]$	0,0359	12	18	15,4
$[13\ ;15]$	$[-1{,}8;-1{,}1]$	0,0998	14	48	42,9
$[15\ ;17]$	$[-1{,}1\ ;-0{,}4]$	0,2089	16	93	89,8
$[17\ ;19]$	$[-0{,}4\ ;0{,}3]$	0,2733	18	104	117,5
$[19\ ;21]$	$[0{,}3\ ;1{,}0]$	0,2234	20	100	96,1
$[21\ ;23]$	$[1{,}0\ ;1{,}7]$	0,1141	22	40	49,1
$[23;+\infty)$	$[1{,}7\ ;+\infty)$	0,0446	24	27	19,2
Total		1		430	430

d)

$$X_5^2 = \frac{(18-15{,}4)^2}{15{,}4} + \frac{(48-42{,}9)^2}{42.9} + \frac{(93-89{,}8)^2}{89{,}8} + \frac{(104-117{,}5)^2}{117{,}5} + \frac{(100-96{,}1)^2}{96{,}1} + \frac{(40-49{,}1)^2}{49{,}1} + \frac{(27-19{,}2)^2}{19{,}2} = 7{,}7$$

Como o limite crítico a 0,95 da tabela do Qui-quadrado é 11,1, conclui-se que a hipótese de normalidade dos dados não deve ser rejeitada.