CAPÍTULO 5

DISTRIBUIÇÕES DE PROBABILIDADE

PPGEP
UFRGS

Introdução

O histograma é usado para apresentar dados amostrais extraídas de uma população.

Por exemplo, os 50 valores de uma característica dimensional apresentados anteriormente representam uma amostra de um processo industrial.

O uso de métodos estatísticos permite que se analise essa amostra e se tire algumas conclusões sobre o processo de manufatura.

Introdução

Uma ***distribuição de probabilidade*** é um modelo matemático que relaciona um certo valor da variável em estudo com a sua probabilidade de ocorrência.

Há dois tipos de distribuição de probabilidade:

1. Distribuições Contínuas: Quando a variável que está sendo medida é expressa em uma escala contínua, como no caso de uma característica dimensional.

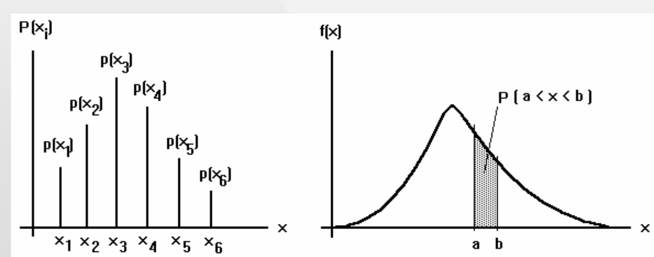
2. Distribuições Discretas: Quando a variável que está sendo medida só pode assumir certos valores, como por exemplo os valores inteiros: 0, 1, 2, etc.

Introdução

No caso de distribuições discretas, a probabilidade de que a variável X assuma um valor específico x_o é dada por: $P(X = x_o) = P(x_o)$

No caso de variáveis contínuas, as probabilidades são especificadas em termos de intervalos, pois a probabilidade associada a um número específico é zero.

$$P(a \le X \le b) = \int_a^b f(x)\,dx$$



Distribuições de Probabilidade

PPGEP

5.1. Distribuições Discretas Mais Importantes

5.1.1. Distribuição Binomial

5.1.2. Distribuição Poisson

Distribuições de Probabilidade

PPGEP

5.1.1. Distribuição Binomial

A distribuição binomial é adequada para descrever situações em que os resultados de uma variável aleatória podem ser agrupados em apenas duas classes ou categorias.

As categorias devem ser mutuamente excludentes, de forma que não haja dúvidas na classificação do resultado da variável nas categorias e coletivamente exaustivas, de forma que não seja possível nenhum outro resultado diferente das categorias.

Por exemplo, um produto manufaturado pode ser classificado como perfeito ou defeituoso, a resposta de um questionário pode ser verdadeira ou falsa, as chamadas telefônicas podem ser locais ou interurbanas.

Distribuições de Probabilidade

PPGEP

5.1.1. Distribuição Binomial

Mesmo variáveis contínuas podem ser divididas em duas categorias, como por exemplo, a velocidade de um automóvel pode ser classificada como dentro ou fora do limite legal.

Geralmente, denomina-se as duas categorias como *sucesso* ou *falha*. Como as duas categorias são mutuamente excludentes e coletivamente exaustivas:

$$P(sucesso) + P(falha) = 1$$

Consequentemente, sabendo-se que, por exemplo, a probabilidade de sucesso é *P(sucesso) = 0,6*, a probabilidade de falha é *P(falha) = 1-0,6 = 0,4*.

Distribuições de Probabilidade

PPGEP

5.1.1. Distribuição Binomial

Condições de aplicação:

- **são feitas *n* repetições do experimento, onde n é uma constante;**
- **há apenas dois resultados possíveis em cada repetição, denominados sucesso e falha**
- **a probabilidade de sucesso *(p)* e de falha *(1- p)* permanecem constante em todas as repetições;**
- **as repetições são independentes, ou seja, o resultado de uma repetição não é influenciado por outros resultados.**

PPGEP

Distribuições de Probabilidade

5.1.1. Distribuição Binomial

Seja um processo composto de uma seqüência de *n* observações independentes com probabilidade de sucesso constante igual a *p*, a distribuição do número de sucessos seguirá o modelo Binomial:

$$P(x) = \binom{n}{x} p^x (1-p)^{n-x} \qquad \mathbf{x = 0,1,....,n}$$

onde $\binom{n}{x}$ representa o número de combinações de *n* objetos tomados *x* de cada vez, calculado como:

$$\binom{n}{x} = \frac{n!}{x!(n-x)!}$$

PPGEP

Distribuições de Probabilidade

5.1.1. Distribuição Binomial

Os parâmetros da distribuição Binomial são *n* e *p*.

A média e a variância são calculadas como:

$\mu = np$

$\sigma^2 = np(1 - p)$

A distribuição Binomial é usada com freqüência no controle de qualidade quando a amostragem é feita sobre uma população infinita ou muito grande.

Nas aplicações de controle da qualidade, *x* em geral representa o número de defeituosos observados em uma amostra de *n* itens.

PPGEP

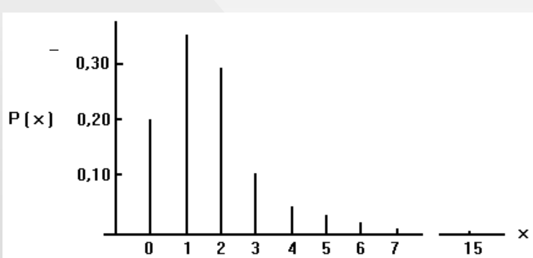
Distribuições de Probabilidade

5.1.1. Distribuição Binomial

Por exemplo, se *p = 0,10* e *n = 15*, a probabilidade de obter *x* itens não conformes é calculada usando a equação da Binomial. Por exemplo, para *x=1*

$$\binom{15}{1} = \frac{15!}{1!(15-1)!} = 15$$

$$\hat{P}(1) = \binom{15}{1} x\, 0{,}10^1\, x\, (1-0{,}10)^{15-1} = 15 \times 0{,}10\, x\, 0{,}23 = 0{,}34$$



PPGEP

Distribuições de Probabilidade

5.1.1. Distribuição Binomial

Outra estatística de interesse para o controle de qualidade é a fração de defeituosos de uma amostra:

$$\hat{p} = \frac{x}{n}$$

$$\sigma^2_{\hat{p}} = \frac{p(1-p)}{n}$$

Distribuições de Probabilidade

PPGEP

5.1.2. Distribuição de Poisson

A distribuição de Poisson é adequada para descrever situações onde existe uma probabilidade de ocorrência em um campo ou intervalo contínuo, geralmente tempo ou área.

Por exemplo, o nº de acidentes por mês, nº de defeitos por metro quadrado, nº de clientes atendidos por hora.

Nota-se que a variável aleatória é discreta (número de ocorrência), no entanto a unidade de medida é contínua (tempo, área).

Além disso, as falhas não são contáveis, pois não é possível contar o número de acidentes que não ocorreram, nem tampouco o número de defeitos que não ocorreram.

Distribuições de Probabilidade

PPGEP

5.1.2. Distribuição de Poisson

Condições de aplicação:

- o número de ocorrências durante qualquer intervalo depende somente da extensão do intervalo;
- as ocorrências ocorrem independentemente, ou seja, um excesso ou falta de ocorrências em algum intervalo não exerce efeito sobre o número de ocorrências em outro intervalo;
- a possibilidade de duas ou mais ocorrências acontecerem em um pequeno intervalo é muito pequena quando comparada à de uma única ocorrência.

Distribuições de Probabilidade

PPGEP

5.1.2. Distribuição de Poisson

A distribuição de Poissson fica completamente caracterizada por um único parâmetro λ que representa a taxa média de ocorrência por unidade de medida.

A equação para calcular a probabilidade de x ocorrências é dada por:

$$P(x) = \frac{e^{-\lambda}\lambda^x}{x!} \qquad x = 0, 1, \ldots$$

A média e a variância da distribuição de Poisson são:

$\mu = \lambda$

$\sigma = \lambda$

Distribuições de Probabilidade

PPGEP

5.1.2. Distribuição de Poisson

A aplicação típica da distribuição de Poisson no controle da qualidade é como um modelo para o número de defeitos (não-conformidades) que ocorre por unidade de produto (por m², por volume ou por tempo, etc.).

A distribuição de Poisson é uma forma limite da distribuição Binomial, quando $n \to \infty$ e $p \to 0$, mas mantendo o quociente $np = \lambda$.

PPGEP

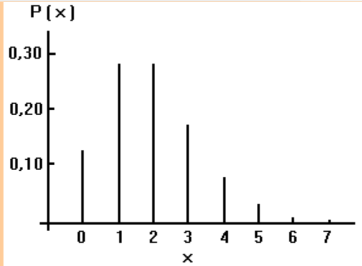
Distribuições de Probabilidade

5.1.2. Distribuição de Poisson

O número de defeitos de pintura segue uma distribuição de Poisson com $\lambda = 2$.

Então, a probabilidade que uma peça apresente mais de 4 defeitos de pintura virá dada por:

$$1 - P\{X \le 4\} = 1 - \sum_{x=0}^{4} \frac{e^{-2} 2^4}{4!} = 1 - 0{,}945 = 0{,}055 = 5{,}5\%$$



x	P(x)
0	0,135
1	0,270
2	0,270
3	0,180
4	0,090
5	0,036
6	0,012

5.2. Distribuições contínuas mais Importantes

- 5.2.1. Distribuição Exponencial
- 5.2.2. Distribuição Weibull
- 5.2.3. Distribuição Normal

5.2.1. Distribuição Exponencial

Na distribuição de Poisson, a variável aleatória é definida como o número de ocorrências em determinado período, sendo a média das ocorrências no período definida como λ.

Na distribuição Exponencial a variável aleatória é definida como o tempo entre duas ocorrências, sendo a média de tempo entre ocorrências de $1/\lambda$.

Por exemplo, se a média de atendimentos no caixa bancário é de λ = 6/min, então o tempo médio entre atendimentos é $1/\lambda$ = 1/6 de minuto ou 10 segundos.

5.2.1. Distribuição Exponencial

Condição de aplicação:

a) o número de ocorrências deve seguir uma distribuição de Poisson.

Se nós considerarmos a distribuição de Poisson como o modelo para o número de ocorrências de um evento no intervalo de *[0,t]* teremos:

$$P(x) = \frac{e^{-\lambda t} (\lambda t)^x}{x!}$$

E nesse caso pode ser demonstrado que a distribuição dos intervalos entre ocorrências irá seguir o modelo Exponencial com parâmetro λ.

5.2.1. Distribuição Exponencial

O modelo da distribuição Exponencial é o seguinte:

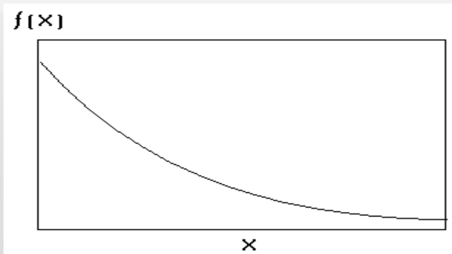
$$f(t) = \lambda e^{-\lambda t}; \qquad t \geq 0$$

onde $\lambda > 0$ é uma constante.

A média e o desvio padrão da distribuição exponencial são calculados usando:

$$\mu = \frac{1}{\lambda}$$

$$\sigma = \frac{1}{\lambda}$$



5.2.1. Distribuição Exponencial

A distribuição Exponencial acumulada vem dada por:

$$F(t) = P\{T \leq t\} = \int_0^t \lambda e^{-\lambda t} dx = 1 - e^{\lambda t} \qquad t \geq 0$$

A distribuição Exponencial é largamente utilizada no campo da confiabilidade, como um modelo para a distribuição dos tempos até a falha de componentes eletrônicos.

Nessas aplicações o parâmetro λ representa a taxa de falha para o componente, e *$1/\lambda$* é o tempo médio até a falha.

5.2.1. Distribuição Exponencial

Por exemplo, suponha que uma máquina falhe em média uma vez a cada dois anos *λ=1/2=0,5*. Calcule a probabilidade da máquina falhar durante o próximo ano.

$$F(t) = P\{T \leq 1\} = 1 - e^{0,5x1} = 1 - 0,607 = 0,393$$

A probabilidade de falhar no próximo ano é de *0,393* e de não falhar no próximo ano é de *1-0,393=0,607*.

Ou seja, se forem vendidos 100 máquinas 39,3% irão falhar no período de um ano.

Conhecendo-se os tempos até a falha de um produto é possível definir os períodos de garantia.

5.2.2. Distribuição de Weibull

O modelo da distribuição de Weibull é:

$$f(x) = \frac{\gamma}{\theta}\left(\frac{x-L}{\theta}\right)^{\gamma-1} e\left[-\left(\frac{x-L}{\theta}\right)^{\gamma}\right]$$

$$F(x) = 1 - e\left[-\left(\frac{x-L}{\theta}\right)^{\gamma}\right]$$

onde:

- γ: **parâmetro de forma**
- θ: **parâmetro de escala**
- *L*: **parâmetro de localização**

Distribuições de Probabilidade

PPGEP

5.2.2. Distribuição de Weibull

A média e a variância da distribuição de Weibull:

$$\mu = L + \theta\,\Gamma\left(1+\frac{1}{\gamma}\right)$$

$$\sigma^2 = \theta^2\left[\Gamma\left(1+\frac{2}{\gamma}\right) - \Gamma\left(1+\frac{1}{\gamma}\right)^2\right]$$

A distribuição de Weibull é muito flexível e pode assumir uma variedade de formas.

Ela tem sido usada extensivamente para modelar tempos de processo ou tempos até a falha de componentes elétricos, componentes mecânicos, elementos estruturais e sistemas complexos.

Distribuições de Probabilidade

PPGEP

5.2.3. Distribuições Normal

A distribuição Normal é a mais importante das distribuições estatísticas, tanto na teoria como na prática:

- Representa a distribuição de freqüência de muitos fenômenos naturais;
- Serve como aproximação da distribuição Binomial, quando ***n*** é grande;
- As médias e as proporções de grandes amostras segue a distribuição Normal (Teorema do Limite Central).

Distribuições de Probabilidade

PPGEP

5.2.3. Distribuições Normal

A distribuição Normal é em forma de sino, unimodal, simétrica em relação à sua média e tende cada vez mais ao eixo horizontal à medida que se afasta da média.

Ou seja, teoricamente os valores da variável aleatória podem variar de -∝ a ∝.

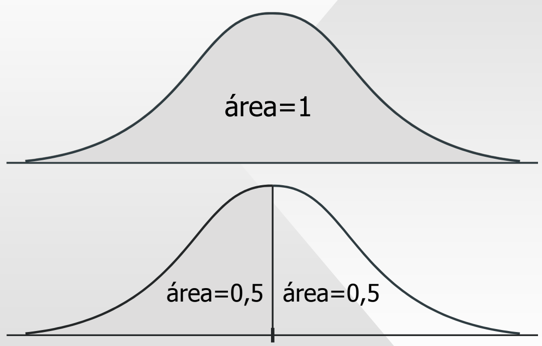
A área abaixo da curva Normal representa 100% de probabilidade associada a uma variável.

A probabilidade de uma variável aleatória tomar um valor entre dois pontos quaisquer é igual à área compreendida entre esses dois pontos.

Distribuições de Probabilidade

PPGEP

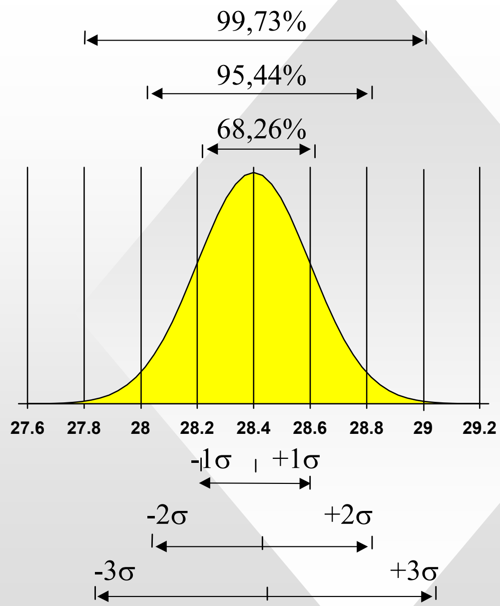
A área total abaixo da curva é considerada como 100%. Isto é, a área total abaixo da curva é 1.



PPGEP

Distribuições de Probabilidade

Percentuais da distribuição Normal:



PPGEP

Distribuições de Probabilidade

5.2.3. Distribuições Normal

A distribuição Normal fica completamente caracterizada por dois parâmetros: a média e o desvio-padrão.

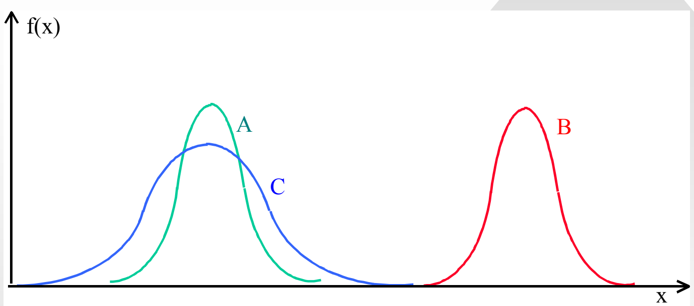
Ou seja, diferentes médias e desvio-padrões originam curvas normais distintas, como se pode visualizar nos exemplos contidos na tabela abaixo onde há amostras provenientes de distribuições com média e desvios-padrões distintos.

Amostras	Dados	Localização ($\bar{x}$)	Variabilidade (R)
A	10 12 14 16 18	$\bar{x} = 14$	$\bar{R} = 8$
B	22 24 26 28 30	$\bar{x} = 26$	$\bar{R} = 8$
C	6 10 14 18 22	$\bar{x} = 14$	$\bar{R} = 16$

PPGEP

Distribuições de Probabilidade

5.2.3. Distribuições Normal



a) da distribuição A para B muda a tendência central, mas a variabilidade é constante;

b) da distribuição A para C muda a variabilidade, mas a tendência central é constante;

c) da distribuição B para C muda a tendência central e a variabilidade.

PPGEP

Distribuições de Probabilidade

5.2.3. Distribuições Normal

Uma conseqüência importante do fato de uma distribuição Normal ser completamente caracterizada por sua média e desvio-padrão é que a área sob a curva entre um ponto qualquer e a média é função somente do número de desvios-padrões que o ponto está distante da média.

Como existem uma infinidade de distribuições normais (uma para cada média e desvio-padrão), transformamos a unidade estudada seja ela qual for (peso, espessura, tempo, etc.) na unidade Z, que indica o número de desvios-padrão a contar da média.

Distribuições de Probabilidade

5.2.3. Distribuições Normal

Dessa forma, o cálculo de probabilidades (área sob a curva) pode ser realizado através de uma distribuição Normal padronizada, onde o parâmetro é a variável reduzida *Z.*

A distribuição Normal pode ser representada por uma equação matemática dada por:

$$f(x) = \frac{1}{\sigma\sqrt{2\Pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Distribuições de Probabilidade

5.2.3. Distribuições Normal

A distribuição Normal acumulada é obtida calculando a probabilidade de *X* ser menor que um dado valor *x*:

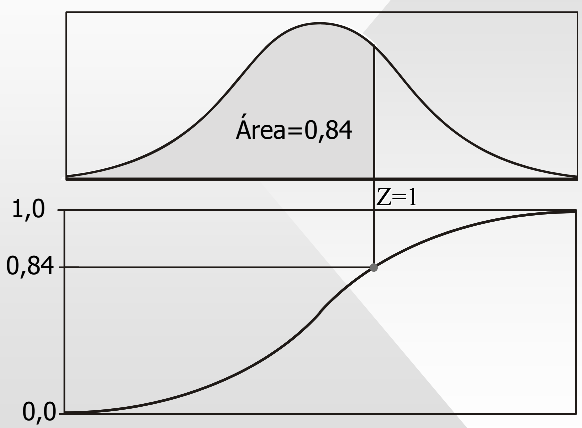
$$P(X \le x) = F(x) = \int_{-\infty}^{x} f(x)\,dx$$

Essa integral não pode ser resolvida em forma fechada, mas a solução está apresentada em tabelas da distribuição Normal padronizada onde se entra com a variável reduzida *Z* (número de desvios-padrões distantes da média) e encontra-se *F(Z)* ou vice-versa.

$$P\{X \le x\} = P\left\{Z \le \frac{x-\mu}{\sigma}\right\} = F(Z) \quad \Rightarrow \text{Tabelado}$$

Distribuições de Probabilidade

Para sabermos o valor da probabilidade, utilizamos a tabela da distribuição Normal. Essa tabela nos fornece a área acumulada até o valor de Z



Distribuições de Probabilidade

As áreas correspondentes as probabilidades da distribuição normal padrão estão tabeladas.

Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
1.0	**0.8413**	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	**0.8770**	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9278	0.9292	0.9306	0.9319

Probabilidade de ocorrência de valores abaixo de Z

5.2.3. Distribuições Normal

O cálculo da variável reduzida *Z* faz uma transformação dos valores reais em valores codificados.

Essa transformação é feita descontando-se a média para eliminar o efeito de localização (tendência central) e dividindo-se pelo desvio-padrão para eliminar o efeito de escala (variabilidade).

Uma vez calculada a variável reduzida *Z*, consulta-se a tabela Normal padronizada para identificar a probabilidade acumulada à esquerda de *Z*,

ou seja, a probabilidade de ocorrerem valores menores ou iguais a um certo valor de *Z* consultado.

5.2.3. Distribuições Normal

Exemplo 1: A resistência à tração do papel usado em sacolas de super-mercado é uma característica de qualidade importante.

Sabe-se que essa resistência segue um modelo Normal com média 40 psi e desvio padrão 2 psi.

Se a especificação estabelece que a resistência deve ser maior que 35 psi, qual a probabilidade que uma sacola produzida com este material satisfaça a especificação?

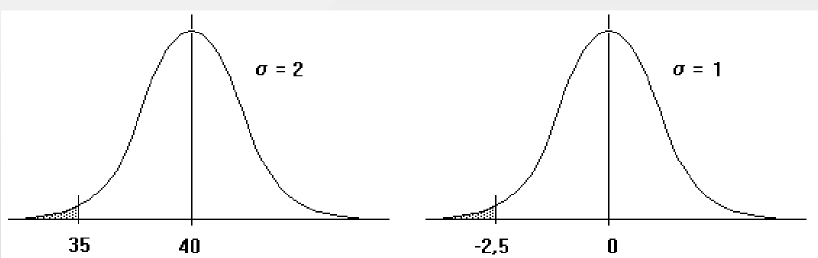
5.2.3. Distribuições Normal

$$P\{X \geq 35\} = 1 - P\{X \leq 35\}$$

$$P\{X \leq 35\} = P\left\{Z \leq \frac{35-40}{2}\right\} = P\{Z \leq -2{,}5\}$$

Tabela: F(-2,5) = 0,0062

Assim a resposta é 1 - 0,0062 = 99,38%



Distribuição para X (valores reais) Distribuição para Z (valores codificados)

5.2.3. Distribuições Normal

Exemplo 2: O diâmetro do eixo principal de um disco rígido segue a distribuição Normal com média 25,08 in e desvio padrão 0,05 in.

Se as especificações para esse eixo são 25,00 ± 0,15 in, determine o percentual de unidades produzidas em conformidades com as especificações.

$$P\{24{,}85 \leq x \leq 25{,}15\} = P\{x \leq 25{,}15\} - P\{x \leq 24{,}85\}$$

$$= P\left\{Z \leq \frac{25{,}15 - 25{,}08}{0{,}05}\right\} - P\left\{Z \leq \frac{24{,}85 - 25{,}08}{0{,}05}\right\}$$

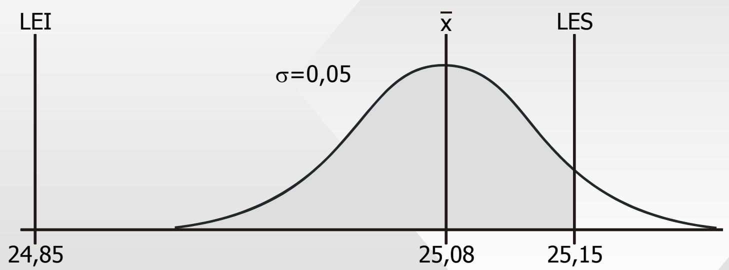
$$= P\{Z \leq 1{,}40\} - P\{Z \leq -4{,}60\} = 0{,}9192 - 0{,}0000 = 0{,}9192$$

ou seja, 91,92% dentro das especificações e 8,08% fora das especificações.

PPGEP

Distribuições de Probabilidade

ou seja, 91,92% dentro das especificações(área cinza) e 8,08% fora das especificações.



PPGEP

Distribuições de Probabilidade

5.2.3. Distribuições Normal

Exemplo 3: No exemplo anterior tem-se cerca de 8% de unidades não-conformes, e essas unidades são invariavelmente do tipo “eixo muito largo”.

Recalcule o percentual de unidades conformes se o processo estivesse centrado em 25,00.

$$P\left\{Z \leq \frac{25{,}15 - 25{,}00}{0{,}05}\right\} - P\left\{Z \leq \frac{24{,}85 - 25{,}00}{0{,}05}\right\}$$

$$P\{Z \leq 3{,}0\} - P\{Z \leq -3{,}0\} = 0{,}9987 - 0{,}00135 = 0{,}9973$$

ou seja, 99,73% dentro das especificações e 0,27% fora das especificações.

PPGEP

Distribuições de Probabilidade

5.2.3. Distribuições Normal

Exemplo 4: Suponha que $X \rightarrow N(85; 9)$. Encontre um valor limite *x*, tal que *P{ X > x}* = 0,05.

$$P\{X > x\} = 1 - P\{X \leq x\} = 1 - P\left\{Z \leq \frac{x-85}{9}\right\} = 0{,}05$$

$$P\left\{Z \leq \frac{x-85}{9}\right\} = 0{,}95 \qquad \text{Tabela } Z = 1{,}645$$

$$\text{Assim,} \quad 1{,}645 = \frac{x-85}{9}; \qquad x = 99{,}805$$

PPGEP

Distribuições de Probabilidade

5.2.3.1. Propriedades da Distribuição Normal

A distribuição Normal tem muitas propriedades úteis.

Uma dessas propriedades é que qualquer combinação linear de variáveis normalmente distribuídas também seguirá o modelo Normal, ou seja:

Se $X_1, X_2, \ldots\ldots\ldots, X_n$ têm distribuição normal e são independentes,

A variável *Y* que é uma combinação linear de *X*:

$Y = a_1X_1 + a_2X_2 + \ldots\ldots + a_kX_k$

5.2.3.1. Propriedades da Distribuição Normal

Também seguirá o modelo normal, com média

$$\mu_Y = a_1\mu_1 + \ldots\ldots + a_n\mu_n$$

e variância

$$\sigma_Y^2 = a_1^2\sigma_1^2 + \ldots\ldots + a_n^2\sigma_n^2$$

onde $a_1, \ldots\ldots, a_n$ são constantes.

Distribuições de Probabilidade

5.2.3.2. Teorema do Limite Central

O Teorema do Limite Central indica que a soma (e por conseguinte a média) de *n* variáveis independentes seguirá o modelo Normal, independentemente da distribuição das variáveis individuais.

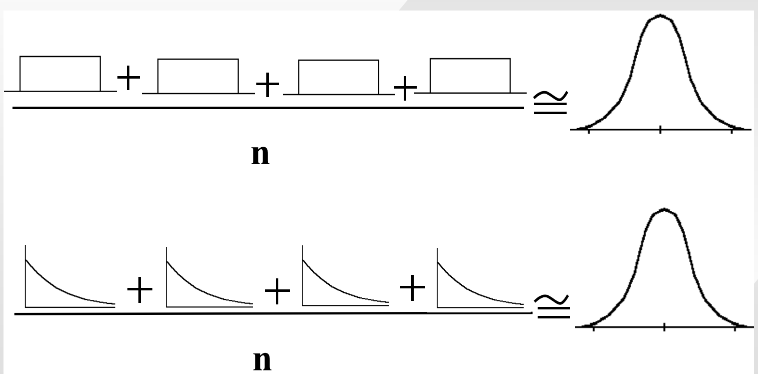
A aproximação melhora na medida em que *n* aumenta. Se as distribuições individuais não são muito diferentes da Normal, basta *n = 4* ou *5* para se obter uma boa aproximação.

Se as distribuições individuais forem radicalmente diferentes da Normal, então será necessário *n = 20* ou mais.

Distribuições de Probabilidade

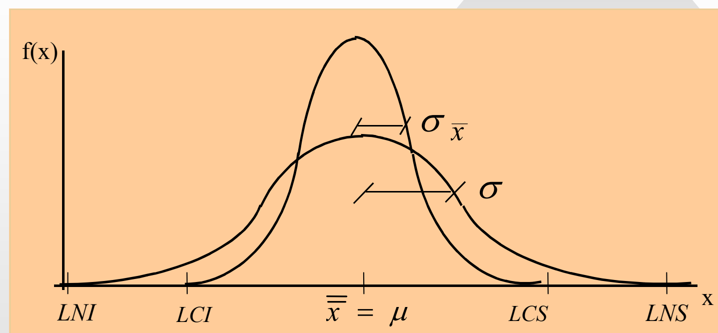
5.2.3.2. Teorema do Limite Central

Na figura abaixo pode ser visto um desenho esquemático do teorema do limite central.



Distribuições de Probabilidade

5.2.3.2. Teorema do Limite Central



Os limites da distribuição dos valores individuais são chamados de limites naturais e os limites da distribuição de probabilidade das médias são chamados de limites de controle.

Distribuições de Probabilidade

PPGEP

Distribuições de Probabilidade

5.2.3.2. Teorema do Limite Central

Exemplo 5: **A distribuição de probabilidade da variável resultante do lançamento de um dado segue a distribuição uniforme, ou seja, qualquer valor (1,2,3,4,5,6) tem a mesma probabilidade (1/6) de ocorrer.**

No entanto, se ao invés de lançar um dado, sejam lançados dois dados e calculada a média, a média dos dois dados seguirá uma distribuição aproximadamente Normal.

PPGEP

Distribuições de Probabilidade

1º dado	2º dado	Soma	Média
1	1	2	1,0
1	2	3	1,5
2	1	3	1,5
1	3	4	2,0
3	1	4	2,0
2	2	4	2,0
1	4	5	2,5
4	1	5	2,5
3	2	5	2,5
2	3	5	2,5
1	5	6	3,0
5	1	6	3,0
2	4	6	3,0
4	2	6	3,0
3	3	6	3,0
1	6	7	3,5
6	1	7	3,5
2	5	7	3,5

1º dado	2º dado	Soma	Média
5	2	7	3,5
3	4	7	3,5
4	3	7	3,5
2	6	8	4,0
6	2	8	4,0
3	5	8	4,0
5	3	8	4,0
4	4	8	4,0
3	6	9	4,5
6	3	9	4,5
4	5	9	4,5
5	4	9	4,5
4	6	10	5,0
6	4	10	5,0
5	5	10	5,0
5	6	11	5,5
6	5	11	5,5
6	6	12	6,0

PPGEP

Distribuições de Probabilidade

5.2.3.2. Teorema do Limite Central

Tabela de freqüência da média dos dois dados

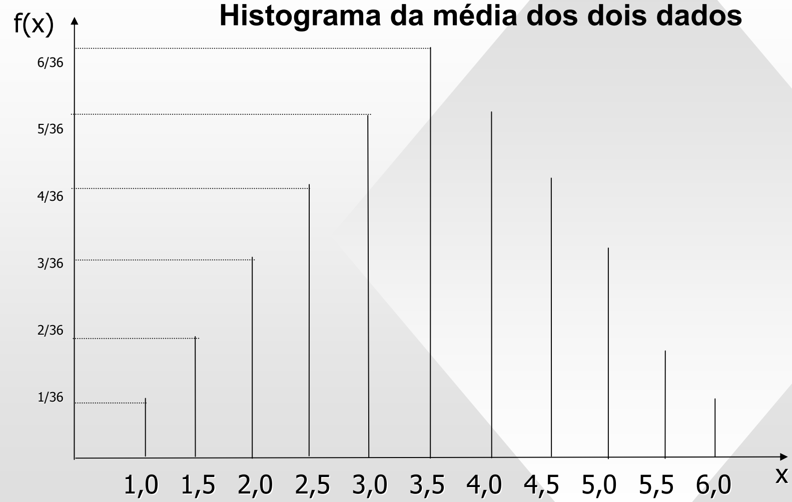
Média de dois dados	Freqüência
1,0	1
1,5	2
2,0	3
2,5	4
3,0	5
3,5	6
4,0	5
4,5	4
5,0	3
5,5	2
6,0	1

PPGEP

Distribuições de Probabilidade

5.2.3.2. Teorema do Limite Central

Histograma da média dos dois dados



Distribuições de Probabilidade

PPGEP

5.2.3.2. Teorema do Limite Central

Conforme pode ser visto no histograma anterior, o histograma da média dos dois dados resulta aproximadamente Normal. Além disso, observa-se que a aproximação da distribuição Normal melhora na medida que se fizesse a média do lançamento de mais dados.

Distribuições de Probabilidade

PPGEP

5.2.3.2. Teorema do Limite Central

O teorema do limite central é básico para a maioria das aplicações do controle estatístico da qualidade.

O CEP trabalha com a média das amostras, pois independente da distribuição dos valores individuais, a média desses valores irá seguir aproximadamente a distribuição Normal.

A partir do teorema do limite central, sabe-se que a distribuição amostral das médias apresenta os seguintes parâmetros:

$$\bar{\bar{X}} = \mu \qquad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

	População	Amostra
Média	μ	$\bar{x}$
Desvio-padrão	σ	S

Distribuições de Probabilidade

PPGEP

Exemplo 1

Um pesquisador deseja saber média da idade dos alunos de pós-graduação. Supondo que a população dos alunos seja:

25	35	24	43	35	22	49	56
34	26	35	52	40	35	35	25
61	42	58	56	45	40	38	45
33	53	22	35	23	25	36	39

$$\mu = \frac{\sum xi}{N} = \frac{25 + ... + 39}{32} = 38{,}19$$

$$\sigma = \sqrt{\frac{(x_i - \mu)^2}{N}} = \sqrt{\frac{(25 - 38{,}19)^2 + ... + (39 - 38{,}19)^2}{32}} = 11{,}11$$

Distribuições de Probabilidade

PPGEP

Supondo que não fosse possível analisar a população inteira, e os dados fossem coletados por amostras de tamanho n=4

	1	2	3	4	5	6	7	8
	25	35	24	43	35	22	49	56
	34	26	35	52	40	35	35	25
	61	42	58	56	45	40	38	45
	33	53	22	35	23	25	36	39
Média ($\bar{x}$)	38,25	39	34,75	46,5	35,75	30,5	39,5	41,25
Desvio (S)	15,69	11,40	16,52	9,40	9,43	8,43	6,45	12,92

$$\bar{\bar{x}} = \frac{\sum \bar{x}_i}{k} = \frac{38{,}25 + ... + 41{,}25}{8} = 38{,}18$$

$$\hat{\sigma}_{\bar{x}} = \sqrt{\frac{\sum\left(\bar{x}_i - \bar{\bar{x}}\right)^2}{k-1}} = \sqrt{\frac{(38{,}25 - 38{,}18)^2 + ... + (41{,}25 - 38{,}18)^2}{8-1}} = 4{,}75$$

$$\bar{\bar{x}} = 38{,}18 \quad \cong \quad \mu = 38{,}19$$

$$\sigma_{\bar{x}} = 4{,}75 \quad \cong \quad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{11{,}11}{\sqrt{4}} = 5{,}55$$

PPGEP

Distribuições de Probabilidade

Supondo que não fosse possível analisar a população inteira, e os dados fossem coletados por amostras de tamanho n=8

	1	2	3	4
	25	24	35	49
	34	35	40	35
	61	58	45	38
	33	22	23	36
	35	43	22	56
	26	52	35	25
	42	56	40	45
	53	35	25	39
Média ($\bar{x}$)	**38,62**	**40,62**	**33,12**	**40,37**
Desvio (S)	**12,71**	**13,94**	**8,74**	**9,50**

$$\bar{\bar{x}} = \frac{\sum \bar{x}_i}{k} = \frac{38{,}62 + \ldots + 40{,}37}{4} = 38{,}18$$

$$\hat{\sigma}_{\bar{x}} = \sqrt{\frac{\sum (\bar{x}_i - \bar{\bar{x}})^2}{k-1}} = \sqrt{\frac{(38{,}62-38{,}18)^2 + \ldots + (40{,}37-38{,}18)^2}{4-1}} = 3{,}49$$

$$\bar{\bar{x}} = 38{,}18 \quad \cong \quad \mu = 38{,}19$$

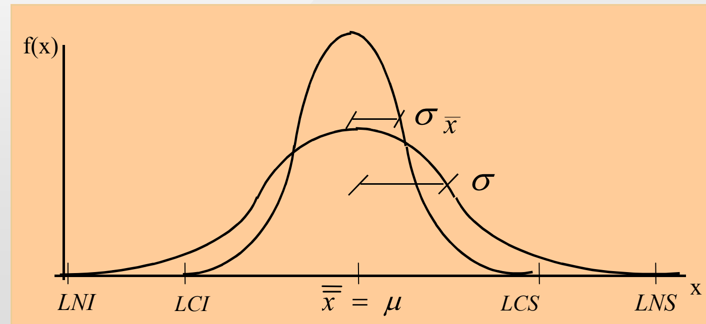
$$\hat{\sigma}_{\bar{x}} = 3{,}49 \cong \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{11{,}11}{\sqrt{8}} = 3{,}93$$

PPGEP

Distribuições de Probabilidade

5.2.3.2. Teorema do Limite Central

A média das médias amostrais é igual a média dos valores individuais e o desvio-padrão das médias é menor do que o desvio-padrão dos valores individuais na razão de $1/\sqrt{n}$



PPGEP

Distribuições de Probabilidade

Exercícios

5.1 Suponha que dois dados sejam lançados e seja *X* a soma dos valores obtidos. Descreva o espaço amostral deste experimento e determine a distribuição de probabilidade de *X*.

5.2 Um processo industrial opera com média de 1% de defeituosos. Baseado em amostras de 100 unidades, calcule as probabilidades de uma amostra apresentar 0, 1, 2, 3 e 4 defeituosos. Plote a distribuição de probabilidade correspondente.

PPGEP

Distribuições de Probabilidade

Exercícios

5.3 Imagine que para o processo anterior, fossem coletadas amostras de 50 unidades e o critério para parar o processo e procurar causas especiais fosse *X=1* ou mais. Calcule a percentagem de vezes que o processo seria interrompido logo após a amostragem.

5.4 Em uma indústria automotiva, defeitos superficiais de pintura ocorrem a uma taxa de 0,15 defeitos / unidade. Encontre a probabilidade de que uma unidade escolhida ao acaso apresente 1 ou mais defeitos superficiais.

Distribuições de Probabilidade

PPGEP

Exercícios

5.5 O setor financeiro de uma loja de departamentos está tentando controlar o número de erros cometido na emissão das notas fiscais. Suponha que esses erros sigam o modelo de Poisson com média $\lambda = 0{,}03$. Qual a probabilidade de uma nota selecionada ao acaso conter 1 ou mais erros?

5.6 A resistência à tração de isoladores cerâmicos apresenta distribuição Normal com média 95 Kg e desvio padrão 4 Kg. Se são produzidas 10.000 unidades desses isoladores, quantos apresentarão resistência inferior a 85 Kg? E quantos apresentarão resistência superior a 90 Kg?

Distribuições de Probabilidade

PPGEP

Exercícios

5.7 A saída de uma bateria segue o modelo Normal com média 12,15 V e desvio padrão 0,2 V. Encontre o percentual que irá falhar em atender às especificações 12 V $\pm$ 0,5 V.

5.8 Se *X* representa medições feitas em um processo que segue o modelo Normal com média 100 e desvio padrão 10, que comportamento irá seguir a média de amostras de 4 unidades retiradas desse processo? E qual será o comportamento da média de 9 unidades retiradas desse processo?

Distribuições de Probabilidade

PPGEP

Exercícios

5.9. Os tempos até a falha de um dispositivo eletrônico seguem o modelo Exponencial, com uma taxa de falha $\lambda = 0{,}012$ falhas/hora. Plote a distribuição de probabilidade correspondente. Depois indique: qual a probabilidade de um dispositivo escolhido ao acaso sobreviver a 50 horas? E a 100 horas?

5.10 O tempo até a venda de um certo modelo de eletrodoméstico, que é regularmente abastecido em um supermercado, segue uma distribuição Exponencial, com parâmetro $\lambda = 0{,}4$ aparelhos/dia. Indique a probabilidade de um aparelho indicado ao acaso ser vendido logo no primeiro dia.

Distribuições de Probabilidade

PPGEP

Exercícios

5.11. Num lote que tem 2% de defeituosos, foram retiradas 40 peças, que será rejeitado se forem encontradas duas ou mais peças defeituosas. Qual a probabilidade de rejeitar o lote?

5.12. Os registros de uma pequena companhia indicam que 40% das faturas por ela emitidas são pagas após o vencimento. De 14 faturas expedidas, determine a probabilidade de:

a)nenhuma ser paga com atraso.

b)no máximo 2 serem pagas com atraso.

c)pelo menos 3 serem pagas com atraso.

d)uma ser paga em dia.

Exercícios

5.13. Uma amostra de 3 m de cabo foi retirada de uma bobina. O cabo tem em média uma falha por m. Qual a probabilidade de não encontrar falha na amostra?

5.14. O tempo necessário, em uma oficina, para o conserto de transmissão para certo carro é normalmente distribuído com média 45 min e desvio padrão 8 min. O mecânico planeja começar o conserto do carro 10 min após o cliente deixá-lo na oficina, comunicando que o carro estará pronto em 1 h. Qual a probabilidade de que o cliente tenha que esperar caso o mecânico esteja enganado e o cliente fique esperando?

Exercícios

5.15 Uma fábrica de pneus fez um teste para medir o desgaste de seis pneus e verificou que ele seguia o comportamento de uma curva normal com média 48.000 km e desvio padrão de 2.000 km. Calcule a probabilidade de um pneu escolhido ao acaso:

a) dure mais que 47.000 km?

b) dure entre 45.000 e 51.000 km?

c) até que quilometragem duram 90% dos pneus?

Exercícios

5.16) O consumo de gasolina por Km rodado para certo tipo de carro, tem distribuição normal com média de 100 ml com desvio padrão de 5 ml.

a) calcular a probabilidade de um carro consumir entre 92 e 106 ml.

b) sabe-se que 73,24% dos carros consumem menos que certa quantidade de gasolina qual é essa quantidade?

c) num grupo de 5 carros qual a probabilidade de dois consumirem mais que 107 ml?

5.17.) Em uma indústria trabalham 1260 pessoas, cujos salários tem média $34.600 e desvio padrão $ 8.500. Calcule a probabilidade de ser inferior a $34.100 o valor da média de uma amostra aleatória constituída por 300 pessoas e 100 pessoas