Inferência Estatística

Aula 5

Agosto de 2008

Mônica Barros

Conteúdo- Revisão de Probabilidade

- As principais distribuições contínuas
  - Distribuição Uniforme
  - Distribuição Exponencial
  - Função Gama
  - Distribuição Gama
  - Relação entre as Densidades Gama e Exponencial
  - Distribuição Qui-quadrado
  - Distribuição Normal
  - Combinações Lineares de Variáveis Normais

Revisão

- É importante relembrar o que caracteriza uma densidade de probabilidade. Seja X uma variável aleatória contínua com densidade de probabilidade f(x) e função de distribuição F(x). Então:

- 1) f(x) ≥ 0 para todo x
- $$2) \int_{-\infty}^{\infty} f(x)dx = 1$$

Revisão

- A probabilidade de X estar num intervalo qualquer é:

$$\Pr(c < X < d) = \int_{c}^{d} f(x)dx$$

- A função de distribuição e a densidade estão relacionadas através de:

$$f(x) = \frac{dF(x)}{dx}$$

Distribuição Uniforme

- Se X ~ Unif(a,b) então sua densidade é:

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

- A função de distribuição é dada por:

$$F(x) = \Pr(X \le x) = \begin{cases} 0 \text{ se } x \le a \\ \dfrac{x-a}{b-a} \text{ se } x \in (a,b) \\ 1 \text{ se } x \ge b \end{cases}$$

- Note que a *função de distribuição* é *linear* no intervalo (a,b).

Distribuição Uniforme

- Média e Variância da distribuição Uniforme
- Se X ~ Unif(a,b) então:

$$E(X) = \frac{a+b}{2}, VAR(X) = \frac{(b-a)^2}{12} \text{ e}$$

$$M(t) = E(e^{tX}) = \frac{e^{tb} - e^{ta}}{t(b-a)} \text{ onde } t \neq 0$$

Distribuição Exponencial

- A distribuição exponencial é útil na descrição de variáveis aleatórias tais como:
  - Tempo entre ocorrências consecutivas de um evento,
  - Tempo necessário para realização de uma tarefa,
  - Tempo entre chegadas consecutivas de carros em um posto de pedágio,
  - Tempo de atendimento em um caixa de banco.
- Também, a distribuição Exponencial é freqüentemente usada para modelar a vida útil ou duração de componentes eletrônicos.

Distribuição Exponencial

- Se X tem densidade Exponencial(λ) então:

$$f(x) = \begin{cases} \lambda . e^{-\lambda . x} \text{ se } x \ge 0 \text{ e } \lambda > 0 \\ 0 \text{ se } x < 0 \end{cases}$$

- A função de distribuição é:

$$F(x) = \Pr(X \le x) = \begin{cases} 0 \text{ se } x < 0 \\ 1 - e^{-\lambda . x} \text{ se } x \ge 0 \end{cases}$$

Distribuição Exponencial

- Média, Variância e fgm
  - Se X é Exponencial com parâmetro λ, então:

$$E(X) = 1/\lambda$$

$$VAR(X) = 1/\lambda^2$$

$$M(t) = E(e^{tX}) = \left(1 - \frac{t}{\lambda}\right)^{-1} = \left(\frac{\lambda}{\lambda - t}\right)$$ que só existe

quando $t < \lambda$

- Parametrização alternativa
  - Escreva a densidade em termos de β = 1/λ, a média da densidade

Função Gama

- Seja α um número real maior que zero, não necessariamente inteiro.
- A função Gama com argumento α é definida por:

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1}.e^{-t}dt$$

- Esta função será usada para definir uma densidade de probabilidade.

Função Gama

Propriedades da Função Gama

1) Γ(n) = (n-1).Γ(n-1) para n > 1

A demonstração deste fato usa integração por partes.

2) Γ(n) = (n-1)! se n é inteiro > 1

3) Γ(1) = 0! = 1

$$4)\ \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$$

Na verdade, o resultado 4) está intimamente ligado à demonstração de que uma variável Normal padrão integra a 1

Distribuição Gama

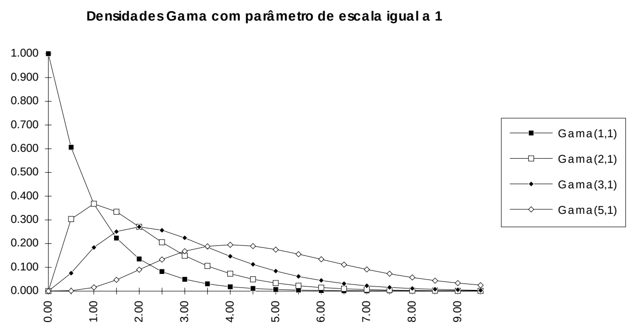
- Seja X uma variável aleatória contínua definida no intervalo (0,∞).
- Dizemos que X tem densidade Gama com parâmetros α e β, e escrevemos X ~ Gama(α, λ) se a densidade de X é:

$$f(x) = \frac{\lambda^{\alpha}.x^{\alpha-1}}{\Gamma(\alpha)}.e^{-\lambda.x}$$

- Os parâmetros α e λ são números reais positivos. α é conhecido como parâmetro de forma, e λ é o parâmetro de escala.

Distribuição Gama

- Algumas densidades Gama(α, λ =1)



Distribuição Gama

- Média, Variância e fgm de uma variável aleatória Gama

- Se X ~ Gama (α, λ) então:

$$E(X) = \frac{\alpha}{\lambda} \;,\; VAR(X) = \frac{\alpha}{\lambda^2} \;\text{ e}$$

$$M(t) = E(e^{tX}) = \left(1 - \frac{t}{\lambda}\right)^{-\alpha} = \left(\frac{\lambda}{\lambda - t}\right)^{\alpha} \;\text{ definida se } t < \lambda$$

Relação entre as Densidades Gama e Exponencial

- Sejam X_1, X_2, X_3, ..., X_k variáveis aleatórias independentes, todas com a mesma densidade Exponencial(λ). Seja $Y = X_1 + X_2 + ... + X_k$. Então Y tem densidade Gama com parâmetros k e λ, onde λ é o parâmetro da Exponencial e k é o número de termos na soma. (Mas por que será que eu me preocupo tanto com a soma de variáveis aleatórias independentes??)

- Além disso, a densidade Exponencial(λ) é apenas um caso particular da densidade Gama quando α = 1 e λ é um número positivo qualquer.

Distribuição Qui-quadrado

- Dizemos que X tem densidade Qui-quadrado com n graus de liberdade, onde n é um número inteiro positivo se a sua densidade é:

$$f(x) = \frac{1}{2^{n/2}.\Gamma\left(\frac{n}{2}\right)}.x^{\frac{n}{2}-1}.e^{-\frac{x}{2}}$$

- Notação: X ~ χ_n^2
- Acima X é uma variável aleatória definida no intervalo (0, ∞).

Distribuição Qui-quadrado

- **A densidade Qui-quadrado é apenas um caso particular da densidade Gama onde α = n/2 e λ = 1/2.**

- **A partir das expressões para a média e variância de uma densidade Gama podemos obter facilmente a *média, a variância e a fgm de uma variável Qui-quadrado*.**

Distribuição Qui-quadrado

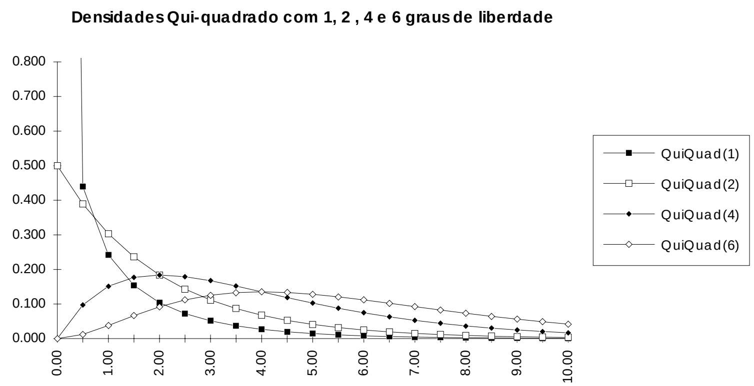
- **Se X é Qui-quadrado com n graus de liberdade então:**

$$E(X) = \frac{n/2}{1/2} = n$$

$$VAR(X) = \frac{n/2}{(1/2)^2} = 2n$$

$$M(t) = E(e^{tX}) = \left(\frac{1}{1-2t}\right)^{n/2} \text{ desde que } t < 1/2$$

Distribuição Qui-Quadrado



Distribuição Qui-quadrado

- **Tabelas da função de distribuição Qui-quadrado**
  - **A densidade Qui-quadrado é tabelada para diversos graus de liberdade.**
  - **As tabelas geralmente fornecem o valor $x_{1-\alpha}$ tal que Pr (X < $x_{1-\alpha}$) = 1 - α para α = 1% , 5%, 10%. Também existem tabelas que apresentam o valor x_α tais que Pr(X < x_α) = α, isto é, Pr(X > x_α) = 1- α.**

Distribuição Qui-quadrado

probabilidade → / graus de liberdade ↓	0.01	0.05	0.10	0.25	0.50	0.75	0.90	0.95	0.99
1	0.000	0.004	0.016	0.102	0.455	1.323	2.706	3.841	6.635
2	0.020	0.103	0.211	0.575	1.386	2.773	4.605	5.991	9.210
3	0.115	0.352	0.584	1.213	2.366	4.108	6.251	7.815	11.345
4	0.297	0.711	1.064	1.923	3.357	5.385	7.779	9.488	13.277
5	0.554	1.145	1.610	2.675	4.351	6.626	9.236	11.070	15.086
6	0.872	1.635	2.204	3.455	5.348	7.841	10.645	12.592	16.812
7	1.239	2.167	2.833	4.255	6.346	9.037	12.017	14.067	18.475
8	1.647	2.733	3.490	5.071	7.344	10.219	13.362	15.507	20.090
9	2.088	3.325	4.168	5.899	8.343	11.389	14.684	16.919	21.666
10	2.558	3.940	4.865	6.737	9.342	12.549	15.987	18.307	23.209
11	3.053	4.575	5.578	7.584	10.341	13.701	17.275	19.675	24.725
12	3.571	5.226	6.304	8.438	11.340	14.845	18.549	21.026	26.217
13	4.107	5.892	7.041	9.299	12.340	15.984	19.812	22.362	27.688
14	4.660	6.571	7.790	10.165	13.339	17.117	21.064	23.685	29.141
15	5.229	7.261	8.547	11.037	14.339	18.245	22.307	24.996	30.578
16	5.812	7.962	9.312	11.912	15.338	19.369	23.542	26.296	32.000
17	6.408	8.672	10.085	12.792	16.338	20.489	24.769	27.587	33.409
18	7.015	9.390	10.865	13.675	17.338	21.605	25.989	28.869	34.805
19	7.633	10.117	11.651	14.562	18.338	22.718	27.204	30.144	36.191
20	8.260	10.851	12.443	15.452	19.337	23.828	28.412	31.410	37.566

Distribuição Qui-quadrado

- Função de Distribuição Qui-quadrado no Excel
- Use as funções DIST.QUI e INV.QUI
- A tabela anterior foi produzida usando INV.QUI – dada uma probabilidade e o grau de liberdade, a função INV.QUI retorna o ponto correspondente da densidade tal que a probabilidade de estar ACIMA do ponto é a especificada como argumento da função.

Distribuição Qui-quadrado

- Função de Distribuição Qui-quadrado no Excel
- Por exemplo, para uma Qui-quadrado com 10 graus de liberdade:
- INV.QUI(0.99, 10) = 2.558
- INV.QUI(0.01,10) = 23.209
- Ou seja, a probabilidade de uma v.a. Qui-quadrado com 10 graus de liberdade exceder 2.558 é 0.99, e a probabilidade da mesma variável exceder 23.209 é 0.01.

Distribuição Qui-quadrado

- A densidade Qui-quadrado é importante no contexto de amostras aleatórias Normais, na estimação da variância amostral.

- Também pode-se provar que o quadrado de uma variável Normal padrão (que estudaremos a seguir) tem densidade Qui-quadrado com um grau de liberdade.

Distribuição Normal

- A distribuição Normal é talvez a mais importante das distribuições de probabilidade.
- Muitos fenômenos físicos ou econômicos são freqüentemente modelados pela distribuição Normal.
- É utilizada para descrever inúmeras aplicações práticas:
  - Altura e peso de pessoas e objetos
  - Nível de chuvas
  - Altura de árvores em uma floresta

Distribuição Normal

- A distribuição Normal tem a forma de um sino, e possui dois parâmetros, μ e σ^2 .
- A distribuição Normal é também chamada de Gaussiana em homenagem ao matemático Carl Friederich Gauss (1777 - 1855).
- A distribuição Normal também funciona como uma boa aproximação para outras densidades. Por exemplo, sob algumas condições pode-se provar que a densidade Binomial pode ser aproximada pela Normal.

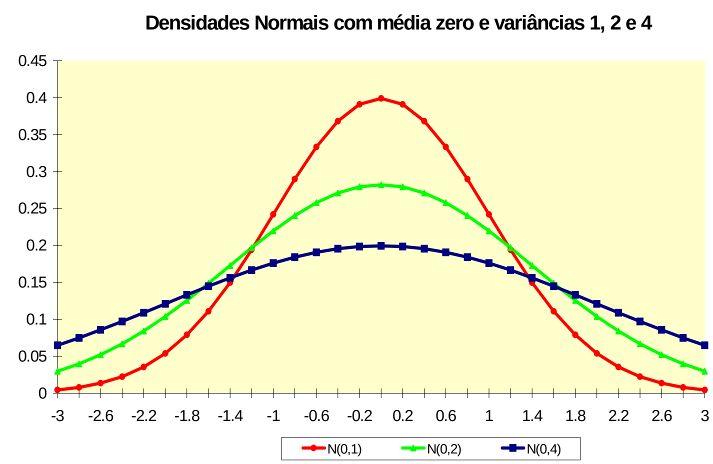
Distribuição Normal

- Densidade Normal com média μ e variância σ^2

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}.e^{-\frac{(x-\mu)^2}{2\sigma^2}} \text{ onde } \sigma^2 > 0 \; e \; \mu \in \mathrm{R}$$

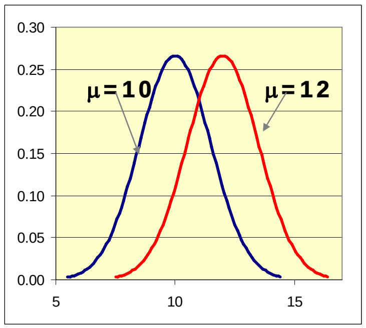
- Notação: X ~ N(μ , σ^2)
- A densidade é simétrica em torno de μ, e quanto maior o valor da variância σ^2, mais "espalhada" é a distribuição.

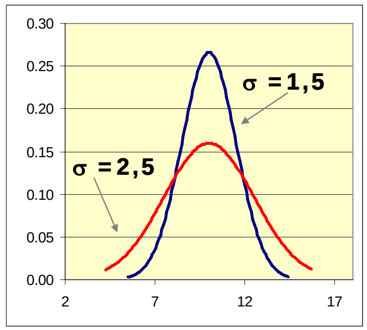
Distribuição Normal



Distribuição Normal

- A distribuição normal é totalmente caracterizada por sua média μ e seu desvio-padrão σ
- A média define o deslocamento horizontal da curva, enquanto o desvio-padrão define o seu achatamento





Distribuição Normal

- **Propriedades**

1) f(x) como definida integra a 1.

2) f(x) > 0 sempre.

3) Os limites de f(x) quando x tende a + ∞ e - ∞ são iguais a zero.

4) A densidade $N(\mu, \sigma^2)$ é *simétrica em torno de μ*, ou seja:

$$f(\mu + x) = f(\mu - x)$$

5) O valor máximo de f(x) ocorre em $x = \mu$

6) Os pontos de inflexão de f(x) são $x = \mu + \sigma$ e $x = \mu - \sigma$.

Distribuição Normal

- **Média, Variância, fgm e função de distribuição**
- **Se X ~ $N(\mu, \sigma^2)$ então:**

$E(X) = \mu$,

$VAR(X) = \sigma^2$

$$M(t) = E(e^{tX}) = \exp\left(\mu.t + \frac{\sigma^2 t^2}{2}\right)$$

- **A sua função de distribuição é:**

$$F(x) = \Pr(X \le x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{\frac{-(u-\mu)^2}{2\sigma^2}\right\} du$$

Distribuição Normal

- **Não é possível resolver analiticamente esta integral - precisamos de uma tabela!**
- **Tabela: será feita para a distribuição N(0,1)**
- É possível transformar uma variável $N(\mu,\sigma^2)$ numa N(0,1) sem grandes dificuldades, e então podemos **tabelar os valores da função de distribuição de uma N(0,1)**, e esta tabela pode ser usada para encontrar probabilidades envolvendo qualquer variável aleatória Normal.

Distribuição Normal

- **Problema:**
  Não é possível criar uma tabela para cada uma das (infinitas) densidades Normais existentes.
- **Solução:**
  Trabalha-se com a densidade Normal com média 0 e variância 1, e converte-se todas as outras Normais para esta, chamada de Normal padrão ou Normal standard.

  A maioria dos livros de estatística fornece tabelas de probabilidade para a distribuição normal padronizada.

Distribuição Normal

- **Transformação numa N(0,1)**
- **Se X ~ N(μ , σ^2) então Z = (X – μ)/σ é uma variável Normal com média 0 e variância 1.**

- **Logo, para transformar uma variável aleatória Normal com quaisquer parâmetros numa Normal (0,1) você deve:**
  1- Subtrair a média
  2- Dividir o resultado por σ, o desvio padrão

  A variável aleatória resultante deste procedimento é uma N(0,1).

Distribuição Normal

- **Se X pertence a uma distribuição normal com média μ e desvio-padrão σ, seu valor normalizado é dado por:**

$$Z = \frac{X - \mu}{\sigma}$$

- **Existem dois tipos de tabela, que fornecem basicamente a mesma coisa:**
  - Pr(0≤ Z ≤ z_0), ou seja, a probabilidade do lado direito da curva normal a partir da média até o valor z_0
  - $\Phi(z_0)$ = Pr (Z ≤ z_0) = 0.5 + Pr (0≤ Z ≤ z_0) (por que?)

- **Iremos trabalhar com a tabela da função de distribuição, isto é: $\Phi(z_0)$**

Distribuição Normal

- **Toda variável Normal pode ser transformada numa Normal com média 0 e variância 1.**

- **Logo, só existe a necessidade de criar uma única tabela para a função de distribuição acumulada.**

- **Se X é N(μ , σ^2) . Então a variável Z = (X - μ) /σ tem distribuição Normal com média zero e variância um, isto é, Z é N(0,1).**

Distribuição Normal

- **Cálculo de probabilidades**

Se X é uma variável Normal com média μ e desvio padrão σ então:

$$\Pr(a \le X \le b) = \Pr\left(\frac{a-\mu}{\sigma} \le \frac{X-\mu}{\sigma} \le \frac{b-\mu}{\sigma}\right) = \Pr\left(\frac{a-\mu}{\sigma} \le Z \le \frac{b-\mu}{\sigma}\right) =$$

$$= \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

- **onde Φ é a função de distribuição da N(0,1), que é tabelada. Alguns valores importantes são:**
Φ(1.645) = 0.95 , Φ (1.96) = 0.975 e Φ (2.326) = 0.99

Distribuição Normal

- **O Excel fornece diretamente o valor de $\Phi(z_0)$ através da função DIST.NORMP.**

- **O único argumento para esta função é o valor z_0 para o qual você quer calcular a probabilidade de estar abaixo, pois a função pressupõe que a distribuição usada é a Normal padrão (média 0 e variância 1).**

Tabela da N(0,1) usando $\Phi(z_0)$

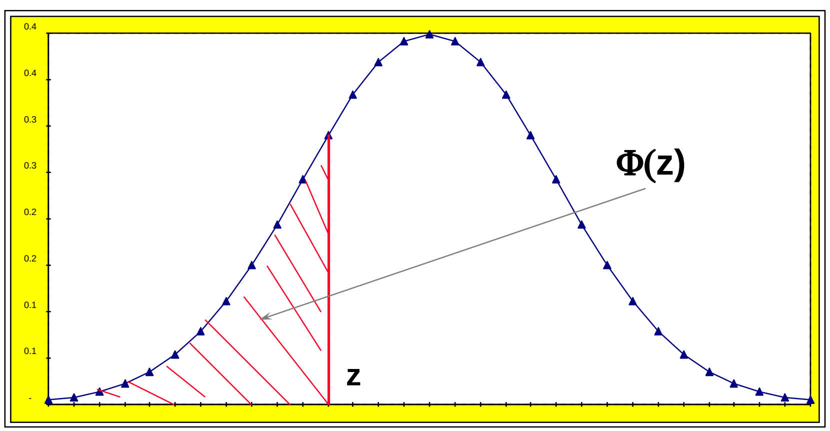


Tabela da N(0,1)

- **Simetrias**
- **Φ(–z) = 1- Φ(z) se z > 0**
- **ISSO É IMPORTANTE POIS A TABELA SÓ CONTÉM VALORES DE z POSITIVOS!**
- **Probabilidade de um intervalo simétrico em torno de zero**
- **Pr (-t < Z < t) = 1 - 2{Φ(-t) } = 1 - 2 {1 - Φ(t)} = 2. Φ(t) - 1 onde Z ~ N(0,1)**

Tabela da N(0,1) ($\Phi(z_0) = Pr(Z \leq z_0)$)

z	ϕ(z)	z	ϕ(z)	z	ϕ(z)	z	ϕ(z)
0.00	0.5000	0.62	0.7324	1.24	0.8925	1.86	0.9686
0.02	0.5080	0.64	0.7389	1.26	0.8962	1.88	0.9699
0.04	0.5160	0.66	0.7454	1.28	0.8997	1.90	0.9713
0.06	0.5239	0.68	0.7517	1.30	0.9032	1.92	0.9726
0.08	0.5319	0.70	0.7580	1.32	0.9066	1.94	0.9738
0.10	0.5398	0.72	0.7642	1.34	0.9099	1.96	0.9750
0.12	0.5478	0.74	0.7704	1.36	0.9131	1.98	0.9761
0.14	0.5557	0.76	0.7764	1.38	0.9162	2.00	0.9772
0.16	0.5636	0.78	0.7823	1.40	0.9192	2.02	0.9783
0.18	0.5714	0.80	0.7881	1.42	0.9222	2.04	0.9793
0.20	0.5793	0.82	0.7939	1.44	0.9251	2.06	0.9803
0.22	0.5871	0.84	0.7995	1.46	0.9279	2.08	0.9812
0.24	0.5948	0.86	0.8051	1.48	0.9306	2.10	0.9821
0.26	0.6026	0.88	0.8106	1.50	0.9332	2.12	0.9830
0.28	0.6103	0.90	0.8159	1.52	0.9357	2.14	0.9838
0.30	0.6179	0.92	0.8212	1.54	0.9382	2.16	0.9846
0.32	0.6255	0.94	0.8264	1.56	0.9406	2.18	0.9854
0.34	0.6331	0.96	0.8315	1.58	0.9429	2.20	0.9861
0.36	0.6406	0.98	0.8365	1.60	0.9452	2.22	0.9868
0.38	0.6480	1.00	0.8413	1.62	0.9474	2.24	0.9875
0.40	0.6554	1.02	0.8461	1.64	0.9495	2.26	0.9881
0.42	0.6628	1.04	0.8508	1.66	0.9515	2.28	0.9887
0.44	0.6700	1.06	0.8554	1.68	0.9535	2.30	0.9893
0.46	0.6772	1.08	0.8599	1.70	0.9554	2.32	0.9898
0.48	0.6844	1.10	0.8643	1.72	0.9573	2.34	0.9904
0.50	0.6915	1.12	0.8686	1.74	0.9591	2.36	0.9909
0.52	0.6985	1.14	0.8729	1.76	0.9608	2.38	0.9913
0.54	0.7054	1.16	0.8770	1.78	0.9625	2.40	0.9918
0.56	0.7123	1.18	0.8810	1.80	0.9641	2.42	0.9922
0.58	0.7190	1.20	0.8849	1.82	0.9656	2.44	0.9927
0.60	0.7257	1.22	0.8888	1.84	0.9671	2.46	0.9931

Tabela da N(0,1)

- Dicas
- Você precisa explorar as simetrias da N(0,1) pois a tabela só é dada para valores positivos de z_0. Por causa da simetria em torno de zero, $\Phi(0) = 0.5$ e $\Phi(z_0)$ é menor que 0.5 se z_0 for um número negativo.
- **Se você tiver dúvidas, faça um desenho!**
- Lembre-se sempre que $\Phi(z_0)$ é uma função de distribuição, ou seja, mede a probabilidade de estarmos ABAIXO do ponto z_0.

Tabela da N(0,1)

- **Dica - uso da calculadoras da série HP48**
- A função UTPN fornece $1 - \Phi(z_0)$ (na verdade, esta função é até mais geral, pois pode ser usada com uma média e variância qualquer; consulte o manual da sua calculadora).
- O algoritmo que irei mostrar é válido apenas para a N(0,1).
  - Menu MTH
  - Submenu PROB
  - Função UTPN
  - Argumentos 0, 1, z_0
  - Retorna a probabilidade de uma variável N(0,1) estar ACIMA do ponto z, ou seja, $1 - \Phi(z_0)$

Distribuição Normal

- Seja $X \sim N(\mu, \sigma^2)$ e $k > 0$. Mostre que $Pr\{ \mu - k\sigma < X < \mu + k\sigma \}$ só depende de k (não depende de μ e σ).
- Solução
- Note que a probabilidade desejada é a probabilidade de X estar a uma distância menor ou igual a k desvios padrões da sua média.

Distribuição Normal

$$\Pr(\mu - k\sigma < X < \mu + k\sigma) = \Pr(-k\sigma < X - \mu < +k\sigma) =$$
$$= \Pr\left(-\frac{k\sigma}{\sigma} < \frac{X-\mu}{\sigma} < +\frac{k\sigma}{\sigma}\right) = \Pr\left(-k < \frac{X-\mu}{\sigma} < k\right) = \Pr(-k < Z < +k) =$$
$$= 2.\Phi(k) - 1$$

- **As probabilidades para alguns valores k estão abaixo:**

Pr(μ - σ < X < μ + σ) = 2.Φ(1) - 1 = 0.6826

Pr(μ - 1.645σ < X < μ + 1.645σ) = 2.Φ(1.645) - 1 = 0.90

Pr(μ - 1.96σ < X < μ + 1.96σ) = 2.Φ(1.96) - 1 = 0.95

Pr(μ - 2.57σ < X < μ + 2.57σ) = 2.Φ(2.57) - 1 = 0.99

Distribuição Normal

- **Exemplo**

Numa agência bancária localizada numa grande cidade brasileira, verificou-se que os clientes pessoa física mantêm, em média, um volume de R$ 4800,00 aplicados no banco.

A dispersão entre os volumes de recursos, medida pelo desvio padrão, é R$ 1600,00. Além disso, pode-se encarar os saldos dos correntistas como independentes entre si e Normalmente distribuídos.

Distribuição Normal

- **O banco pretende abrir uma nova agência e seus executivos imaginam que o poder aquisitivo nesta nova área é semelhante ao dos clientes desta agência.**

- **a) Um cliente é VIP se está entre os 5% com maior volume de recursos. Quanto uma pessoa deveria manter no banco para ser considerada cliente VIP?**

Distribuição Normal

- **b) O banco pretende cobrar tarifas mais altas dos clientes que têm um baixo volume de recursos aplicados na instituição.**

Os clientes cujos volumes de recursos estão entre os 10% mais baixos terão de pagar esta tarifa mais alta. Abaixo de qual volume um cliente será alvo desta tarifa diferenciada?

Distribuição Normal

- **Solução**

  Seja X a variável que mede o volume de recursos de um cliente típico da agência. Então X é Normal (4800, (1600)²). Daí: $Z = \dfrac{X - 4800}{1600}$

  tem densidade Normal padrão.

  Para estar entre os 5% mais "ricos", precisamos encontrar z_0 tal que $\Phi(z_0)$ = 95%. Usando a função INV.NORMP do Excel, encontramos z_0 = 1.645. Logo,

$$\frac{X - 4800}{1600} = 1.645 \Rightarrow X = 4800 + 1.645(1600) = 7432$$

Distribuição Normal

- **Solução (continuação)**

  b) Para estar entre os 10% mais "pobres" precisamos encontrar z_0 tal que $\Phi(z_0)$ = 10%. A função INV.NORMP do Excel fornece z_0 = -1.281. Logo,

$$\frac{X - 4800}{1600} = -1.281 \Rightarrow X = 4800 - 1.281(1600) = 2750.40$$

- **Ou seja, clientes com volume de recursos abaixo de R$ 2750 estarão sujeitos a uma tarifa mais alta, e aqueles com volume de aplicações acima de R$ 7432 terão tratamento VIP.**

Distribuição Normal

- **Exemplo**
- **O saldo devedor dos usuários de um certo cartão de crédito é uma variável aleatória Normal com média R$ 200 e desvio padrão R$ 75.**
  - **a) Qual a probabilidade do saldo devedor de um usuário estar entre R$ 100 e R$ 300?**
  - **b) Qual deve ser o seu saldo devedor para que você esteja entre os 5% mais endividados?**
- **Solução**

  X é Normal com média 200 e desvio padrão 75 e assim

  Z =(X- 200)/75 é N(0,1).

Distribuição Normal

- **Solução (continuação)**

  Pr(100 < X < 300) =

$$\Pr\left(\frac{100-200}{75} < Z < \frac{300-200}{75}\right) = \Pr(-1.333 < Z < +1.333) =$$

$$= \Phi(1.333) - \Phi(-1.333) = 2.\Phi(1.333) - 1 = 0.8176$$

  b) Para que você esteja entre os 5% mais endividados, o saldo devedor padronizado deve ser igual a 1.645 (veja tabela da Normal). Daí:

$$Z = \frac{X - 200}{75} = 1.645 \Rightarrow X = 200 + 1.645(75) = 323.38$$

  é o saldo para estar entre os 5% com maior saldo devedor.

Distribuição Normal

- Exemplo (para casa)
  - O consumo médio residencial de energia elétrica nos meses de verão numa certa cidade é uma variável Normal com média 210 kWh e desvio padrão 18 kWh.

    a) Qual a probabilidade de que o consumo no verão exceda 225 kWh?

    b) Calcule a probabilidade de que o consumo no verão seja inferior a 190 kWh.

    c) Quanto você deve consumir para estar entre os 2.5% que mais gastam energia?

Distribuição Normal

- Exemplo (para casa)
- Numa certa empresa de informática, o salário *anual* médio dos funcionários com menos de 5 anos de experiência é R$ 24000, com desvio padrão de R$ 3000. Suponha que os salários têm distribuição Normal e calcule os valores pedidos a seguir.

Distribuição Normal

- a) Qual a probabilidade do salário anual de um funcionário qualquer com menos de 5 anos de experiência ser menor que R$ 20000?
- b) Qual deve ser o valor do salário anual de um funcionário com menos de 5 anos de experiência se 95% dos funcionários (com menos de 5 anos de experiência) tem salário abaixo dele?

Distribuição Normal

- c)Toma-se uma amostra de 36 funcionários com menos de 5 anos de experiência. Qual a probabilidade do salário médio na amostra exceder R$ 24500?
- d)Toma-se uma amostra de 12 funcionários com menos de 5 anos de experiência. Qual a probabilidade do maior salário na amostra exceder R$ 28000?

Combinações Lineares de Variáveis Normais

- Sejam X_1, X_2, ..., X_n variáveis aleatórias **independentes**, onde $X_i \sim N(\mu_i, \sigma_i^2)$ e seja $Y = X_1 + X_2 + ... + X_n$.
- Então Y tem distribuição Normal com média μ_y e variância σ_y^2 dadas por:

$$\mu_y = \sum_{i=1}^{n} \mu_i$$

$$\sigma_y^2 = \sum_{i=1}^{n} \sigma_i^2$$

Combinações Lineares de Variáveis Normais

- **Um caso particular importante é: se os X_i´s forem iid $N(\mu, \sigma^2)$, então sua soma é Normal com média $n.\ \mu$ e variância $n.\ \sigma^2$ e a média amostral é Normal com média μ e variância σ^2/n.**

$$X_1, \ldots, X_n \text{ iid } N(\mu, \sigma^2) \Rightarrow \sum_{i=1}^{n} X_i \sim N\left(n\mu, n\sigma^2\right)$$

$$X_1, \ldots, X_n \text{ iid } N(\mu, \sigma^2) \Rightarrow \bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

Distribuição Normal

- **Exemplo (continuação)**
- **Considere o exemplo dos saldos em aplicações bancárias. Suponha que tomamos uma amostra de 16 clientes da agência.**
- **Qual a probabilidade de que o saldo médio das aplicações dos clientes na amostra exceda R$ 4900?**

Seja $\bar{X}$ a média dos saldos dos clientes na amostra.

$\bar{X}$ tem distribuição $N\left(4800, \frac{(1600)^2}{16}\right)$

Distribuição Normal

- **Então:**

$$\Pr(\bar{X} > 4900) = \Pr\left(\frac{\bar{X} - 4800}{1600/4} > \frac{4900 - 4800}{1600/4}\right) = \Pr\left(Z > \frac{100}{400}\right) =$$

$$= \Pr(Z > 0.25) = 1 - \Phi(0.25) = 1 - 0.599 = 0.401$$

Exemplo (para casa)

- Um estudante universitário gasta em média R$ 600,00 em livros por ano. A dispersão entre os valores gastos, medida pelo desvio padrão, é R$ 240,00. Além disso, pode-se encarar os valores gastos pelos universitários como independentes entre si e Normalmente distribuídos. Além disso, a maioria dos estudantes adquire livros pela Internet.

Exemplo (para casa)

- a) Uma grande livraria na Internet pretende oferecer um cartão VIP aos clientes que mais compram livros. Apenas os 1% que mais consomem livros num período de um ano receberão o cartão. Acima de qual volume anual de compras um consumidor se candidata ao cartão VIP?
- b) Considere 16 estudantes universitários. Qual a probabilidade do gasto médio anual em livros destas 16 pessoas ultrapassar R$ 660,00?
- c) Dentre as 16 pessoas nesta mesma amostra, qual a probabilidade do estudante que menos consumiu livros ter gasto mais de R$ 650 no ano?