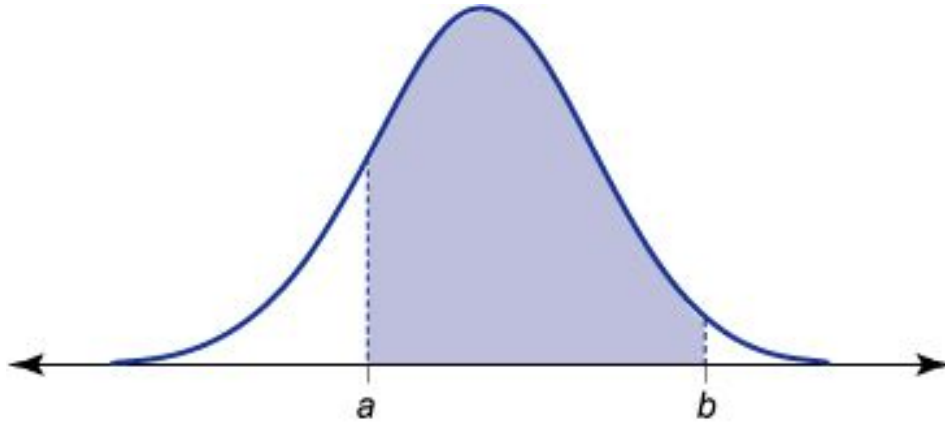


Centro Universitário
Christus- UNICHRISTUS

Estatística Computacional

Especialização em Ciência
de Dados e Inteligência
de Negócios (Big Data e
BI)

Prof. Dr. Manoel Ribeiro



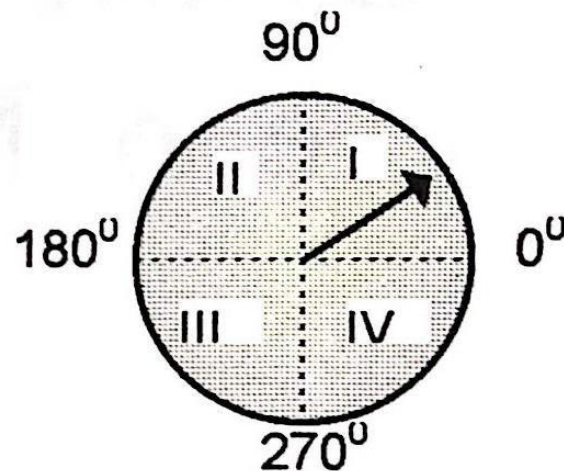
*Principais distribuições
contínuas de
probabilidade*

Variável aleatória contínua

- Representa características medidas num intervalo
 - De números reais
- Exemplo
 - a altura de um indivíduo tomado ao acaso.
 - Pode assumir infinitos valores num intervalo de 0 a 2,5m
- Não se define a probabilidade de um ponto
 - Ex.: 1,682333...
- As probabilidades são definidas para intervalos
 - Ex.: 1,60 a 1,80m

Exemplo de distribuição contínua

EXEMPLO 8.1 Considere um círculo, com medidas de ângulos, em graus, a partir de uma determinada origem, como mostra a figura ao lado. Neste círculo tem um ponteiro que é colocado a girar no sentido anti-horário.



Seja X a variável aleatória que indica o ponto em que o ponteiro pára de girar. Como existem infinitos pontos no intervalo de 0 a 360° , esta variável é contínua. Vejamos, inicialmente, a probabilidade do ponteiro parar no quadrante I, isto é, a probabilidade de X assumir um valor entre 0 e 90° .

Exemplo de distribuição contínua

Admitindo que não exista alguma região de preferência para o ponteiro parar, podemos deduzir, pelo *princípio da equiprobabilidade*, que as probabilidades de parada são iguais para os quatro quadrantes. Assim, a probabilidade do ponteiro parar no primeiro quadrante deve ser igual a $1/4$.

Podemos representar o evento *ponteiro parar no quadrante I* por $0 \leq X < 90$. E esta probabilidade por $P(0 \leq X < 90)$. Em termos de variáveis aleatórias contínuas, os sinais “<” e “ \leq ” são equivalentes, pois, considerando a equiprobabilidade de todos os pontos e, considerando a existência de infinitos pontos, podemos definir a probabilidade de ocorrência de um particular ponto como nula.

A distribuição de probabilidades de uma variável aleatória contínua pode ser representada por uma certa função não negativa, com a área formada entre o eixo das abscissas e a curva desta função, igual a 1 (um). Os eventos podem ser representados por intervalos no eixo das abscissas (eixo X), enquanto as correspondentes probabilidades, por áreas sob a curva. Apresentamos, na Figura 8.1, uma distribuição de probabilidades para o experimento do Exemplo 8.1, sob forma gráfica.

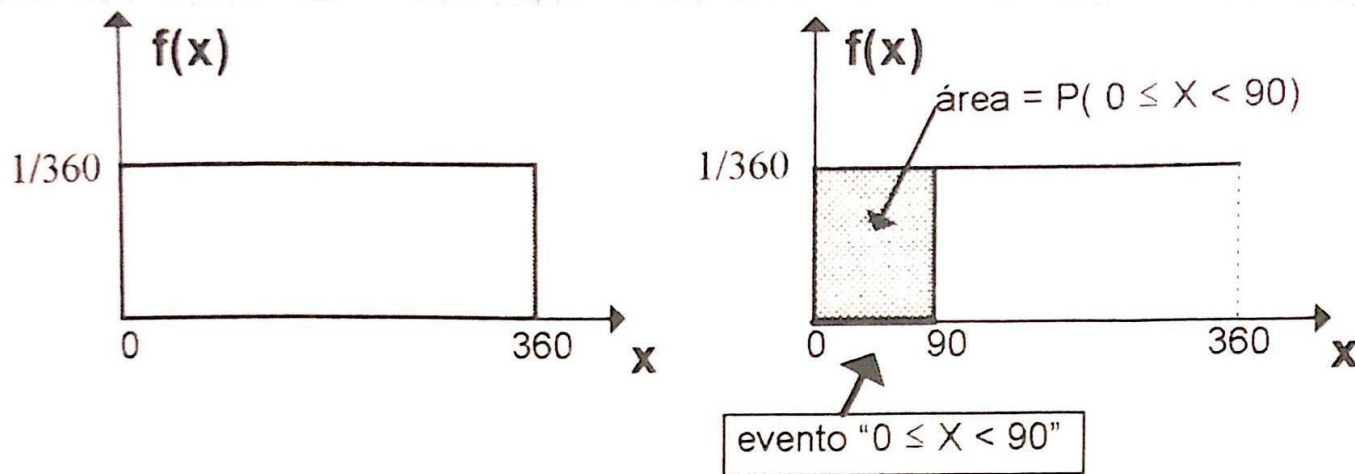


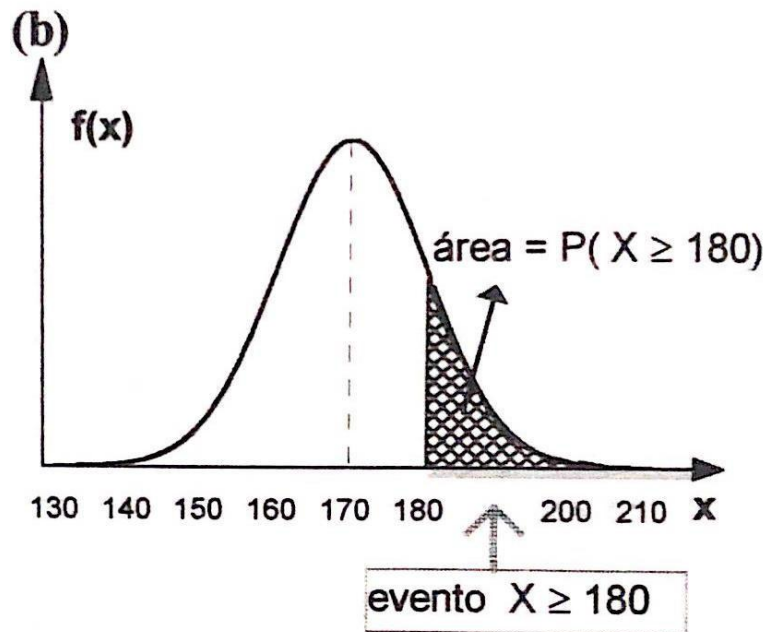
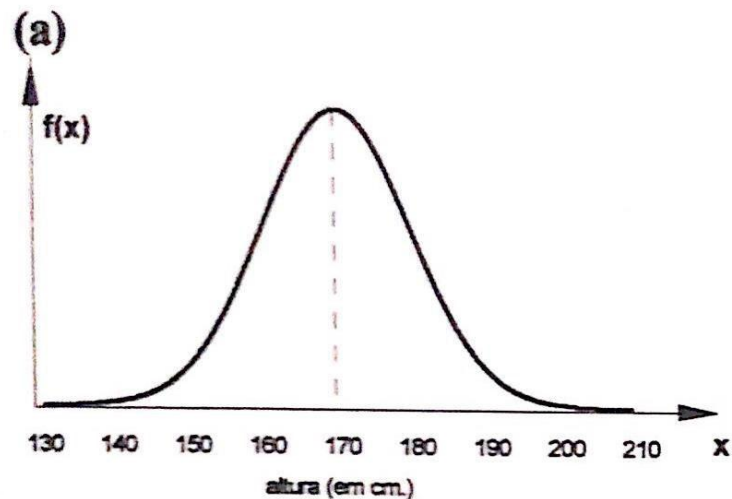
FIG. 8.1 Ilustração de (a) uma distribuição de probabilidades para a variável aleatória do Exemplo 8.1 e (b) a probabilidade do evento " $0 \leq X < 90$ ".

EXEMPLO 8.2 Selecionar, aleatoriamente, de uma certa universidade, um estudante do sexo masculino. Seja X o valor de sua altura, em centímetros.

Temos, novamente, uma variável aleatória contínua, mas, desta vez, não é razoável atribuir a mesma probabilidade para diferentes faixas de altura. Por exemplo, é intuitivo que a probabilidade do estudante acusar altura no intervalo de 165 a 175 cm é bem maior do que no intervalo de 190 a 200 cm, mesmo que ambos os intervalos tenham a mesma amplitude.

A Figura 8.2(a) sugere um modelo mais adequado para a presente situação. Por este modelo, conhecido como *distribuição normal de probabilidades*, existe um *valor típico*, ou *valor médio*, que no caso de alturas de

homens adultos, deve estar em torno de 170 cm. Intervalos em torno deste valor médio têm altas probabilidades de ocorrência, mas as probabilidades diminuem na medida em que nos afastamos deste valor médio, indiferentemente se do lado esquerdo (para valores menores) ou do lado direito (para valores maiores). A Figura 8.2(b) identifica a probabilidade do evento *o estudante sorteado ter mais de 180 cm*.



Distribuição Normal

- Numa distribuição Normal (curva caracterizada por μ e σ)
 - 68,2% das medições estão dentro de 1 σ da média
 - 95,4% estão dentro de 2 σ
 - 99,7% estão dentro de 3 σ
- Útil quando você desconhece
 - Os valores envolvidos (contexto)



Exemplos da Distribuição

A Fig. 8.4 mostra diferentes modelos normais, em termos dos parâmetros μ e σ . Estes modelos podem representar, por exemplo, a distribuição de alturas de crianças, em diferentes populações.

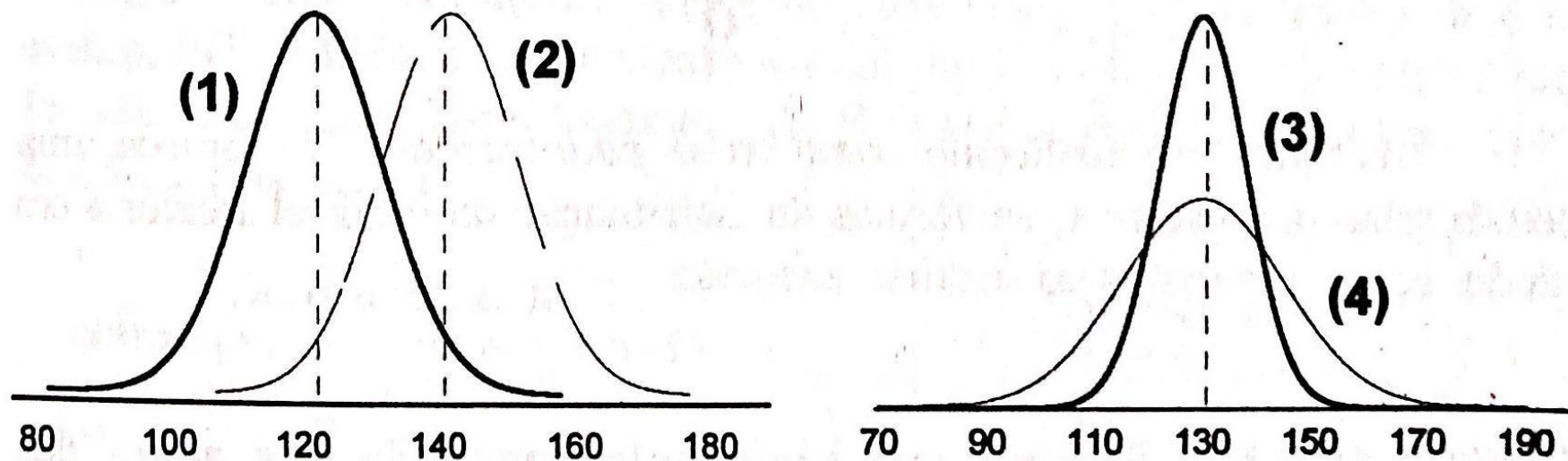


FIG. 8.4 Distribuições normais em função dos parâmetros μ e σ .

Padronizar uma distribuição normal

- Cada domínio tem um intervalo de valores
 - Teríamos que construir uma tabela de consulta
 - Para cada domínio
 - IDEB, PISA, peso, altura, idade...
- Solução alternativa
 - Converter todas para um intervalo padrão
 - Ou seja, uma distribuição normal de $\mu=0$ e $\sigma=1$

Fórmula para padronizar uma distribuição normal

- Para um dado valor x (numa distribuição com média μ e σ)
 - A seguinte fórmula converte os valores x nos valores padronizados z

$$z = \frac{x - \mu}{\sigma}$$

Exemplo de padronização de distribuição

EXEMPLO 8.3 Suponha que numa certa universidade, a altura dos estudantes do sexo masculino tenha distribuição normal com média $\mu = 170$ cm e desvio padrão $\sigma = 10$ cm. A Fig. 8.5 mostra a relação entre a escala dos valores das alturas de universitários masculinos (x) e seus correspondentes valores padronizados (z). Por exemplo, para um estudante de altura $x = 180$ cm, temos o valor padronizado $z = (180 - 170)/10 = 1$, ou seja, este estudante encontra-se a 1 (um) desvio padrão acima da altura média dos estudantes do sexo masculino da universidade.

Exemplo de padronização de distribuição

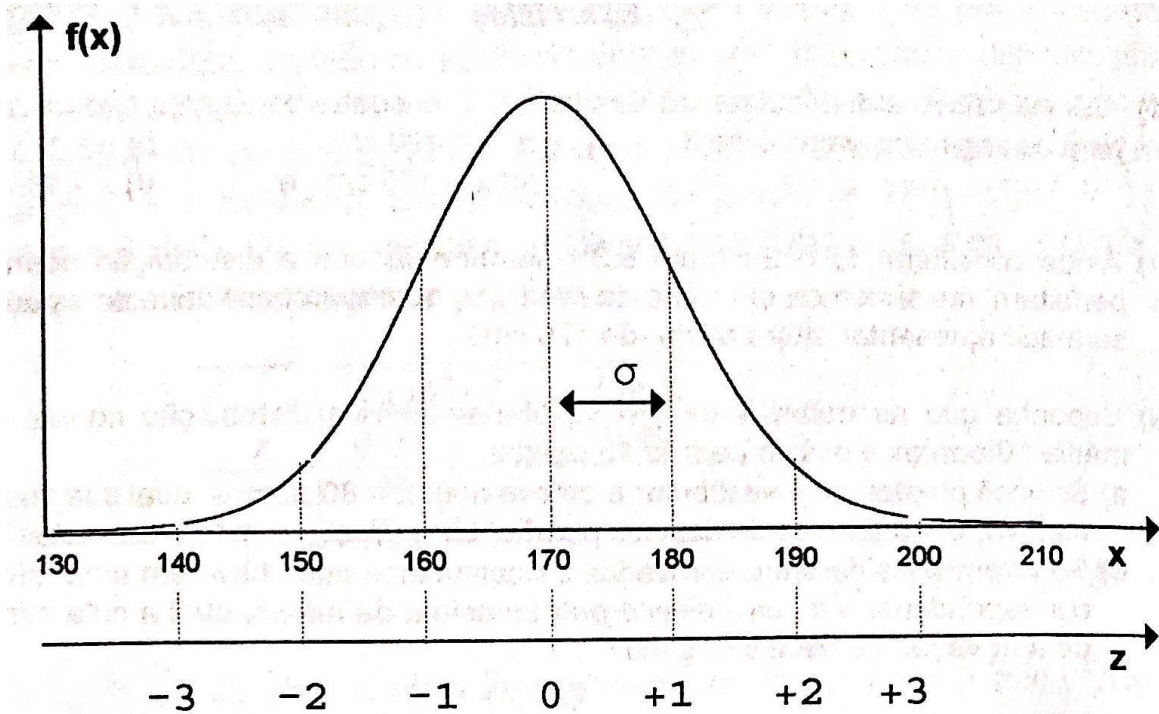


FIG. 8.5 Transformação de valores de alturas de universitários (x) em valores padronizados (z).

Seja X a altura, em centímetro, de um estudante do sexo masculino, selecionado ao acaso, desta universidade. Considere que temos interesse no evento $X > 180$. A Figura 8.6 mostra a equivalência da probabilidade deste evento, $P(X > 180)$, com uma certa área na distribuição normal padrão. Para facilitar a notação, identificaremos por Z uma variável aleatória com distribuição normal padrão.

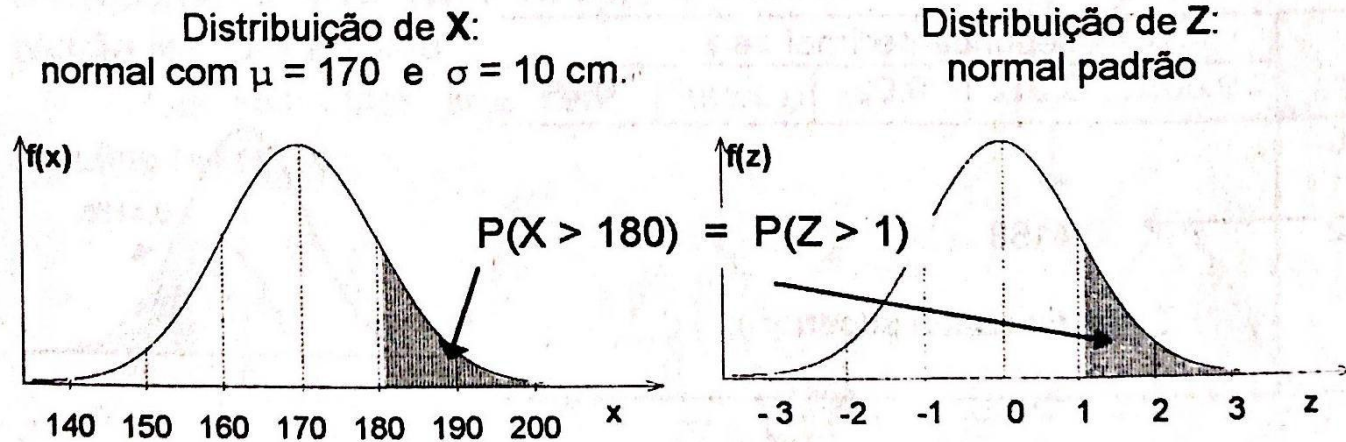


FIG. 8.6 Transformação de um evento da distribuição normal de parâmetros $\mu = 170$ cm e $\sigma = 10$ cm, num evento da distribuição normal padrão.

Tabela da distribuição normal padrão

z	segunda decimal de z				
	0,00	0,01	0,02	...	0,09
0,0					
0,1					
0,2		↓			
...					
		0,4168			
		(área na cauda superior)			

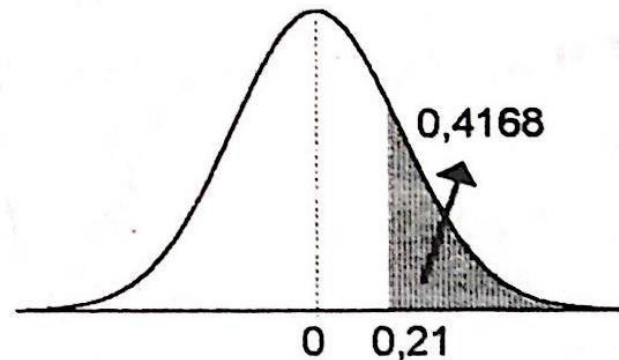


FIG. 8.7 Ilustração do uso da tabela da distribuição normal padrão (Tabela IV do apêndice) para encontrar a área na cauda superior relativa ao valor de $z = 0,21$.

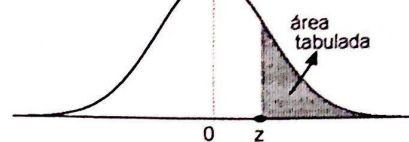
EXEMPLO 8.3, continuação - Admitimos que a altura X de um estudante do sexo masculino, tomado ao acaso de uma universidade, tinha distribuição normal com média 170 cm e desvio padrão 10 cm. Vimos, também, que a probabilidade de ele acusar altura superior a 180 cm, correspondia à área acima de $z = 1$ da curva normal padrão, isto é, $P(X > 180) = P(Z > 1)$. Usando a Tabela IV do apêndice, podemos encontrar esta área (probabilidade), como ilustra o esquema seguinte.

z	segunda decimal de z		
	0,00	...	0,09
...	<div style="text-align: center;">↓</div> 0,1587		
1,0 →			
...			

Portanto:

$$P(X > 180) = 0,1587.$$

TABELA IV- Distribuição normal padrão



z	segunda decimal de z									
	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,4960	0,4920	0,4880	0,4840	0,4801	0,4761	0,4721	0,4681	0,4641
0,1	0,4602	0,4562	0,4522	0,4483	0,4443	0,4404	0,4364	0,4325	0,4286	0,4247
0,2	0,4207	0,4168	0,4129	0,4090	0,4052	0,4013	0,3974	0,3936	0,3897	0,3859
0,3	0,3821	0,3783	0,3745	0,3707	0,3669	0,3632	0,3594	0,3557	0,3520	0,3483
0,4	0,3446	0,3409	0,3372	0,3336	0,3300	0,3264	0,3228	0,3192	0,3156	0,3121
0,5	0,3085	0,3050	0,3015	0,2981	0,2946	0,2912	0,2877	0,2842	0,2843	0,2776
0,6	0,2743	0,2709	0,2676	0,2643	0,2611	0,2578	0,2546	0,2514	0,2483	0,2451
0,7	0,2420	0,2389	0,2358	0,2327	0,2296	0,2266	0,2236	0,2206	0,2177	0,2148
0,8	0,2119	0,2090	0,2061	0,2033	0,2005	0,1977	0,1949	0,1922	0,1894	0,1867
0,9	0,1841	0,1814	0,1788	0,1762	0,1736	0,1711	0,1685	0,1660	0,1635	0,1611
1,0	0,1587	0,1562	0,1539	0,1515	0,1492	0,1469	0,1446	0,1423	0,1401	0,1379
1,1	0,1357	0,1335	0,1314	0,1292	0,1271	0,1251	0,1230	0,1210	0,1190	0,1170
1,2	0,1151	0,1131	0,1112	0,1093	0,1075	0,1056	0,1038	0,1020	0,1003	0,0985
1,3	0,0968	0,0951	0,0934	0,0918	0,0901	0,0885	0,0869	0,0853	0,0838	0,0823
1,4	0,0808	0,0793	0,0778	0,0764	0,0749	0,0735	0,0722	0,0708	0,0694	0,0681
1,5	0,0668	0,0655	0,0643	0,0630	0,0618	0,0606	0,0594	0,0582	0,0571	0,0559
1,6	0,0548	0,0537	0,0526	0,0516	0,0505	0,0495	0,0485	0,0475	0,0465	0,0455
1,7	0,0446	0,0436	0,0427	0,0418	0,0409	0,0401	0,0392	0,0384	0,0375	0,0367
1,8	0,0359	0,0352	0,0344	0,0336	0,0329	0,0322	0,0314	0,0307	0,0301	0,0294
1,9	0,0287	0,0281	0,0274	0,0268	0,0262	0,0256	0,0250	0,0244	0,0239	0,0233
2,0	0,0228	0,0222	0,0217	0,0212	0,0207	0,0202	0,0197	0,0192	0,0188	0,0183
2,1	0,0179	0,0174	0,0170	0,0166	0,0162	0,0158	0,0154	0,0150	0,0146	0,0143
2,2	0,0139	0,0136	0,0132	0,0129	0,0125	0,0122	0,0119	0,0116	0,0113	0,0110
2,3	0,0107	0,0104	0,0102	0,0099	0,0096	0,0094	0,0091	0,0089	0,0087	0,0084
2,4	0,0082	0,0080	0,0078	0,0075	0,0073	0,0071	0,0069	0,0068	0,0066	0,0064

Exercício de Avaliação – 6 (Para Casa)

Seja Z uma variável aleatória com distribuição normal padrão. Use a tabela IV para encontrar as seguintes probabilidades:

- a) $P(Z < 0,42)$
- b) $P(Z < -0,42)$
- c) $P(-0,42 < Z < 0,42)$

Instruções:

- 1) Entregar a solução (valor e a explicação) .

Cálculo de probabilidade com o R

- Cálculo de probabilidade de uma distribuição normal

- `dnorm(x, mean = 0, sd = 1)`

- Retorna o valor pontual

- `pnorm(x, mean = 0, sd = 1, lower.tail = TRUE)`

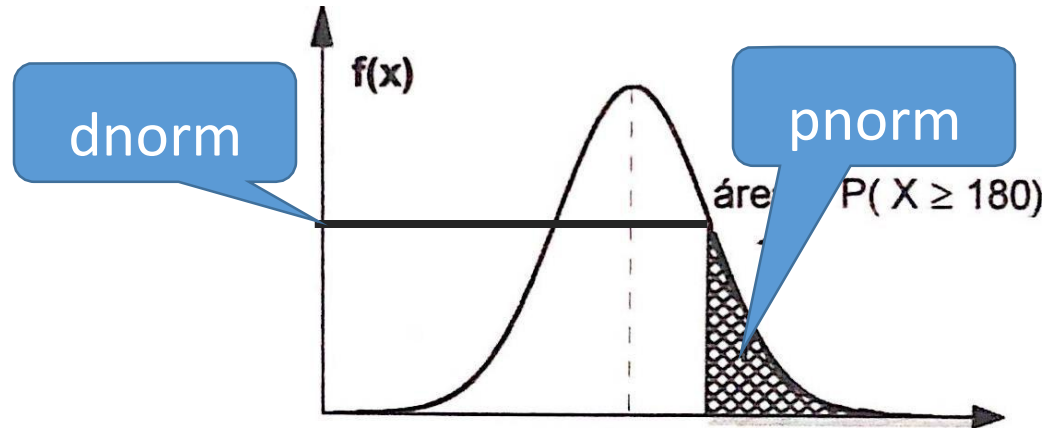
- Retorna a área sob a curva (probabilidade)

- `lower.tail=TRUE`

- Cauda inferior

- `lower.tail=FALSE`

- Cauda superior



Exercício de Avaliação – 7 (Para Casa)

Resolver os itens a, b e c do exercício 6, mas desta vez utilizando o R.

Exercício de Avaliação – 8 (Na Aula)

- Para que você possa entender a utilização da distribuição normal, vamos considerar a situação em que se estudou a durabilidade de um certo tipo de pneu. Verificou-se que esta durabilidade seguia uma distribuição normal com duração média 60.000 km e desvio-padrão 10.000 km. Procurou-se, então, responder os seguintes questionamentos:
 - a) Qual a probabilidade de um pneu aleatoriamente escolhido durar mais de 75.000 km?
 - b) Qual a probabilidade de um pneu aleatoriamente escolhido durar entre 50.000 e 70.000 km?

Fim