UFV - Universidade Federal de Viçosa DPI - Departamento de Informática Prof. André Gustavo dos Santos
INF 222 - Computac¸˜ao Experimental - 2025/2

Exercício 5
Para sábado, 11/out

Distribuição normal, intervalos de confiança e testes de hipótese

---

Mostre as expressões usadas nos cálculos!

1. Seja Z uma variável aleatória normal padrão. Calcule:

a) $\mathbf{P}\{Z < 1.35\}$ b) $\mathbf{P}\{Z \leq 1.35\}$ c) $\mathbf{P}\{Z > 1.35\}$ d) $\mathbf{P}\{|Z| < 1.35\}$ e) $\mathbf{P}\{Z < 5\}$
f) o valor que Z não ultrapassa com probabilidade 0.8

a) Pela tabela de distribuição normal, P{Z < 1.35} = 0.9115
b) Pela tabela de distribuição normal e como há continuidade na variável aleatória então P{Z <= 1.35} = P{Z < 1.35} = 0.9115.
c) P{Z > 1.35} = 1 - P{Z <= 1.35} = 1 - 0.9115 = 0.0885
d) p{z > 1.35} = p{ -1.35 > z > 1.35} = P(Z < 1.35) − P(Z < −1.35), por simetria, P(Z < −1.35) = P(Z > 1.35)
P{Z < -1.35} = 1 - P{Z <= 1.35} = 1 - 0.9115 = 0.0885.
Então, P{|Z| < 1.35} = P{Z < 1.35} - P{Z < -1.35} = 0.9115 - 0.0885 = 0.8230
e) Como a distribuição normal se estende ao infinito, mas a maior parte da probabilidade está concentrada em torno da média, um valor como 5 está muito longe da média. Z = 5 é extremamente grande.
Pela tabela, P(Z<5) ~= 0.9999997
f) Devemos encontrar na tabela de distribuição um valor z tal que o limite da prob. acumulava é até 0.8, esse valor é Z = 0.84.

2. As mesmas perguntas da questão anterior, mas para uma distribuição normal com μ = 1 e σ = 1.5.

Padronizaremos cada item utilizando a fórmula Z = (X - μ) / σ

a) Z = (1.35 - 1)/1.5 ~= 0.2333. Utilizando a tabela, P(Z < 0.23) = 0.5910

Logo,

P(X < 1.35) ~= 0.5910

b) Assim como o item b da questão 1, para uma variável aleatória contínua, a probabilidade de ser menor ou igual a um valor é a mesma que ser estritamente menor que o valor.

Logo,

P(X <= 1.35) ~= 0.5910

c) Pela regra do complemento, P{X > 1.35} = 1 - P{X ≤ 1.35} = 1 - 0.5910 = 0.4090
d) P(−1.35 < X < 1.35); seja Z1 = (-1.35 -1)/1.5 = -1.5667

Z2 = (1.35 -1)/1.5 = 0.2333

Da tabela,

P(Z<0.23) = 0.5910

P(Z < -1.57) = 1 - P(Z < 1.57) = 1 - 0.9418 = 0.0582

Logo,

P(|X| < 1.35) = 0.5910 − 0.0582 = 0.5328

e) Z = (5-1)/1.5 = 2.6667; P(Z <2.6667) = 0.9962
f) Devemos encontrar o valor z pela tabela tal que P(z <= 0.8), pela tabela, \ = 0.84.

Usando a fórmula X = μ + Z * σ = 1 + 0.84 * 1.5 = 2.26, então, X = 2.26

3. Considere a mesma situação de um dos exemplos mostrados em aula, em que supomos a renda familiar mensal média no estado de R$900, com desvio padrão de R$200, seguindo uma distribuição normal.

a) Se o apoio de bolsa alimentação fosse dado a 5% das famílias, sendo beneficiadas as de menor renda familiar mensal, esse apoio seria dado para famílias com até´e que renda?
b) E se o apoio fosse dado às famílias com renda até R$600, quantos % das famílias teriam o apoio?

a) Queremos um valor x tal que P(X <=x) = 0.05. Devemos encontrar um valor Z correspondente a prob. acumulada de 0.05,
pela tabela, P(Z <= -1.645) ~= 0.05. logo Z = -1.645 para esse exemplo,
X = μ + Z * σ = 900 + (-1.645) * 200 = 571
Portanto, o apoio seria dado a famílias com renda de até R$571.

b) Queremos P(X <= 600). Padronizando, Z = (600 - 900)/200 = -1.5
Como,
P(Z <= -1.5) ~= 0.0668,
Portanto,
cerca de 6.68% das famílias receberiam o apoio.

4. A instalação de certo hardware gasta um tempo aleatório com desvio padrão de 5 minutos.

a) Um técnico instalou este hardware em 60 computadores diferentes, com média de tempo de instalação de 42 minutos. Calcule um intervalo para a média real do tempo de instalação com 95% de confiança.
b) Note que o intervalo tem uma margem de erro acima de 1 minuto. Para obter uma margem de até 1 min, a média do tempo de instalação deveria ser obtida a partir de quantas instalações diferentes?

a) Sejam,
n=60, x_ = 42, σ = 5, para 95% de confiança valor usual Z - 1.96
Então,
ME = z* σ/√n = 1.96 * 5/√60 ~= 1.264
Logo,
IC = 42 +- 1.263 => [40.736, 43.264] minutos

b) Queremos que a margem de erro seja no máx. 1 minuto,
Logo,

1 >=z* σ/√n, então, n = (z* σ)^2 = 96.04;

Portanto,

n >= 97, precisaremos de no mínimo 97 instalações.

5. Para assegurar um uso eficiente de um servidor, é necessário estimar seu nu´mero médio de usuários simultâneos . De acordo com dados históricos, obtidos em 100 horários escolhidos ao acaso, o número de usuários simultâneos de certo servidor é 37.4, com desvio padrão 9.1.

a) Indique um intervalo de 90% de confiança¸a para o nu´mero esperado de usuários simultâneos

b) Os dados históricos apresentam evidências com nível de significância de 1% de que o nu´mero médio de usuários simultâneos não passa de 40?

a) Sejam,

n = 100, x_ = 37.4 e s = 9.1

Note que n é grande podemos utilizar a distribuição Z com aproximação, para 90% de confiança Z ~= 1.645

ME = t0,95,99 * s/√n ~= 1.66 * 9.1/10 = 1.66 * 0.91 ~= 1.511

Logo, 37.4 +- 1.511 => [35.889 , 38.911]

b) Sejam,

Hipótese nula H0: μ = 40

Hipótese alternativa H1: μ < 40

α = 0.01

Z crítico para α = 0.01 sendo -2.33 pela tabela

Calculando o valor Z,

Z = (x_ - μ) / (s/√n) = (37.4 - 40) / (9.1 / √100) = -2.6 / (9.1 / 10) ~= -2.857

Comparando Z atual com Z crítico,

-2.857 < -2.33, o valor Z calculado está na região crítica.

Conclusão,

Rejeitamos a hipótese H0, há evidência estatística, ao nível de significância de 1%, de que a média de usuários simultâneos é menor que 40

Mostre as expressões e os dados usados nos cálculos!

6. Os dados abaixo são as alturas, em *cm*, dos presidentes dos EUA no Século XX: 170

178 182 180 183 178 182 188 175 179 185 192 182 183 177 185 188 189

a) Encontre a média, mediana, variância e desvio-padrão das alturas dos presidentes

b) Construa um intervalo de 95% de confiança para a média populacional

c) Teste a hipótese de que a altura dos presidentes é maior que a da média da população (175,3 cm), com nível de significância de 0,05.

a) Seja n = 18,

média:

x_ = soma tam presidentes / 18 = 3256/18 = 180.89cm

mediana:

Ordenando, 170, 175, 177, 178, 178, 179, 180, 182, 182, 182, 183, 183, 185, 185, 188, 188, 189, 192 => (182 + 182) / 2 = 182 cm.

variância:

s^2 = Σ(xi - x_)^2 / (n - 1)

Como,

Σ(xi - x_) = 526.7

s^2 = 526.7 / (18 - 1) = 526.7 / 17 = 30.98 cm^2

Desvio padrão:

s = sqrt(s^2) = sqrt(30.98) ~= 5.57cm

b) Note que o desvio padrão é conhecido e n = 18 é pequeno, logo usaremos a distribuição de Student,

x_ = t a/2 * (s/ sqrt(n)) = 2.110 * 1.313 ~= 2.770

intervalo de confiança,

180.89 +- 2.770 => [178.12, 183.66] cm

c) Sejam,

Hipótese nula H0: μ = 175.3 cm

Hipótese alternativa H1: μ > 175.3 cm (teste unicaudal à direita)

Nível de significância α= 0.05

Graus de liberdade df = 17

O valor t crítico para α= 0.05 e df = 17 é dado por 1.74

Logo,

t = (xx_ - μ) / (s / sqrt(n)) = (180.89 - 175.3) / (5.57 / $\sqrt{}$18) = 5.59 / 1.313 ~= 4.257

Comparando os valores,

4.257 > 1.740, o valor t calculado está na região crítica.

Conclusão,

Rejeitamos a hipótese nula H0. Há evidências com nível de significância de 0.05 de que a altura dos presidentes dos EUA no Século XX é maior que a média da população de 175.3 cm.

7. Refaça o exercício 1 incluindo na lista os presidentes do século XXI: George W. Bush, Barack Obama, Donald Trump e Joe Biden. Há alguma mudança significativa no resultado?

Sejam,

George W. Bush: 182 cm

Barack Obama: 185 cm

Donald Trump: 191 cm

Joe Biden: 183 cm

Então temos,

170, 175, 177, 178, 178, 179, 180, 182, 182, 182, 182, 183, 183, 183, 185, 185, 188, 188, 189, 191, 192

tamanho da amostra = n = 22

a) Seja n = 22,

média:

x_ = soma tam presidentes / 22 = 3997/22~~= 181.62 cm

mediana:

Ordenando, 170, 175, 177, 178, 178, 179, 180, 182, 182, 182, 182, 183, 183, 183, 185, 185, 188, 188, 189, 191, 192 => (182 + 183) / 2 = 182.5 cm.

variância:

s^2 = Σ(xi - x_)^2 / (n - 1)

Como,

Σ(xi - x_) = 654.67

s^2 = 654.67 / (22 - 1) = 31.17 cm^2

Desvio padrão:

s = sqrt(s^2) = sqrt(31.17) ~= 5.58 cm

Conclusão,

Não houve uma mudança significativa nos resultados das medidas de tendência central e dispersão.

8. Teste a hipótese para os presidentes do Brasil em relação à população brasileira. Neste caso, a amostra pode ser menor, visto que a altura não está facilmente disponível para muitos dos presidentes.

Sejam,

média altura população brasileira = µ = 173 cm

desvio padrão da altura população brasileira = σ = 7 cm

Lula 170 cm

Temer 170 cm

Bolsonaro 185 cm

FHC 178 cm

Collor 180 cm

Sarney 182 cm

Castello Branco 173 cm

Médici 170 cm

Geisel 175 cm

Figueiredo 175 cm

tamanho amostra = n = 9

170, 170, 185, 178, 180, 182, 173, 170, 175

Calculando média,

x_ = soma tam presidentes / 9 = 1583/9 ~=175.89 cm

Teste de Hipótese: A altura dos presidentes do Brasil é diferente da média da população brasileira?
Hipótese nula H0: µ = 173 cm (A média das alturas dos presidentes é igual a da população)
Hipótese alternativa H1: µ != 173 cm (A média das alturas dos presidentes é diferente da da população)
Nível de significância α = 0.05
Graus de liberdade df = 8

Logo,
valor t crítico = +- 2.306 para esses valores de α e df
t = (x_ - µ) / (s / sqrt(n)) = (175.89 - 173) / (5.88 / 3) = 2.89 / (5.88 / 3) ~= 1.474

Comparando valores,
-2.306 < 1.474 < 2.306, o valor t calculado não está na região crítica.

Conclusão,
Não rejeitamos a hipótese nula H0. Com base nesta amostra e neste nível de significância (0.05), não há evidências suficientes para concluir que a altura média dos presidentes do Brasil seja significativamente diferente da média da população brasileira de 173 cm.