Testes de Hipótese

Exemplo

Para decidirmos se os computadores de um lote possuem desempenho do tipo **A** ou do tipo **B**, iremos proceder do seguinte modo:

(i) selecionamos uma amostra de 100 computadores, e determinamos o desempenho médio deles;

(ii) se esse desempenho médio for superior a 176 (unidades de desempenho), diremos que são do tipo **B**; caso contrário, do tipo **A**.

Os parâmetros são:
A: $\mu = 175$ e $\sigma = 10$; **B**: $\mu = 177$ e $\sigma = 10$.

Testes de Hipótese

Exemplo

Defina: *Erro do tipo I* – dizer que os computadores são do tipo **B** quando, na realidade, são do tipo **A**.
Erro do tipo II – dizer que são do tipo **A** quando são do tipo **B**.

(a) Qual a probabilidade de se cometer o erro de tipo I? E do erro de tipo II?

(b) Qual deve ser a regra de decisão se fixarmos a probabilidade do erro de tipo I em 5%? Qual a probabilidade do erro de tipo II, nesse caso?

Testes de Hipótese

(a) Note que H_0 : computadores são do tipo A, ou simplesmente $H_0 : \mu = 175$. Nossa região crítica é dada por $RC = \{\bar{X} > 176\}$, isto é, a região em que rejeitamos H_0.

A probabilidade de se cometer o erro do tipo I é igual a probabilidade de rejeitarmos H_0 quando ela é verdadeira. Lembre-se de que $\text{Var}(\bar{X})$ é σ^2/n. Então

$$P(\text{erro tipo I}) = P(\bar{X} > 176 \,|\, \mu = 175) =$$

$$P\left(Z > \sqrt{100}\left(\frac{176 - 175}{10}\right) \,\middle|\, \mu = 175\right) = 1 - \Phi(1) = 0.159$$

(a) De modo análogo, a probabilidade de se cometer o erro de tipo II é a probabilidade de não rejeitarmos H_0 quando ela é falsa. Ela é dada por

$$P(\text{erro tipo II}) = P(\bar{X} \leq 176 \mid \mu = 177) =$$

$$P\left(Z \leq \sqrt{100}\left(\frac{176-177}{10}\right) \middle| \mu = 177\right) = \Phi(-1) = 0.159$$

Testes de Hipótese

(b) Queremos fixar c para que $P(\bar{X} > c|\mu = 175) = 0.05$. Para isto, basta que

$$P(\bar{X} > c|\mu = 175) = P\left(Z > \sqrt{100}\left(\frac{c - 175}{10}\right)\right)$$
$$= 1 - \Phi(c - 175) = 0.05$$

Note que $\Phi(z) = 0.95 \Leftrightarrow z = 1.64$ (obtemos esse valor consultando a tabela da distribuição normal padrão), então temos que o erro do tipo I será igual a 0.05 se $(c - 175) = 1.64$ ou $c = 176.64$.

Testes de Hipótese

(b) Nossa nova regra de decisão agora é classificar o computador como do tipo B se seu desempenho for superior a 176.64. Essa regra tem a probabilidade de se cometer o erro do tipo I fixada em 0.05. Mas, no entanto, temos agora uma probabilidade de se cometer o erro do tipo II de

$$P(\bar{X} \leq 176.64 | \mu = 177) = \Phi\left(\sqrt{100}\frac{176.64 - 177}{10}\right) = 0.359$$

Exemplo

O tempo de execução de um algoritmo (de uma certa empresa) em horas é de 60 minutos com desvio padrão de 20 . Um novo algoritmo fora projetado e, posteriormente, foi tomada uma amostra de nove execuçõess desse algoritmo e medido o tempo de execução, resultando em 50 minutos. Você diria, ao nível de 5%, que há evidência de melhorias?

Testes de Hipótese

Queremos testar a hipótese de que μ, o tempo de execução do novo algoritmo, em horas, é igual ao do algoritmo anterior vs aquele é menor do que este, ou seja, $H_0 : \mu = 60$ vs. $H_1 : \mu < 60$.

Como σ^2 é conhecido, então a estatística do teste é dada por

$$T = \sqrt{n}\left(\frac{\bar{X} - \mu_0}{\sigma}\right)$$

Note que sob H_0, $\bar{X} \sim N(60, 400/9)$, e a distribuição da estatística do teste é portanto $N(0,1)$. A região crítica é $T < c$.

Testes de Hipótese

Sob um nível de significância $\alpha = 0.05$, temos que a hipótese será rejeitada se

$$T = 3(\bar{X} - 60)/20 < c.$$

Para a normal padrão, $P(Z < c) = 0.05 \Leftrightarrow c = -1.64$. Então a região crítica do teste é $3(\bar{X} - 60)/20 < -1.64$, ou simplesmente $\bar{X} < 49.06$.

Como a média observada $\bar{x} = 50$ é superior a 49.06, não rejeitamos a hipótese nula, a um nível de significância de 5% . Ou seja, não há evidências a favor da hipótese de diminuição do tempo de execução. O p-valor, nesse caso (3*(50-60)/20 = -1,5), é dado por $p = P(T < -1,5|H_0) = \Phi(-1,5) = 0,0668$, $(T \sim N(0,1))$. Como $p = 0,0668 > \alpha$, não rejeitamos H_0.

Testes de Hipótese

Exemplo

O administrador de um certo sistema computacional afirma que 20% dos usuários utilizam uma quantidade de memória maior do que o permitido. Para confirmar sua conjectura, ele usou uma amostra de tamanho 50 (usuários), onde 27% procediam dessa forma. Mostre como os usuários poderiam refutar a afirmação. Utilize um nível de significância de 10%.

Testes de Hipótese

O administrador quer testar as hipóteses $H_0 : p = 0.2$ vs $H_1 : p > 0.2$. A região crítica é, portanto, da forma $\hat{p} > c$. A estatística do teste é

$$T = \sqrt{n}\left(\frac{\hat{p} - p_0}{\sqrt{p_0(1 - p_0)}}\right).$$

Note agora que para grandes amostras, a estatística tem distribuição aproximadamente ormal padrão.

Testes de Hipótese

Sob H_0, c é dado por

$$1 - \Phi\left(\sqrt{50}\frac{c - 0.2}{\sqrt{0.2 \cdot 0.8}}\right) = 0.1$$

O valor de z tal que $1 - \Phi(z) = 0.1$ é $z = 1.28$. Então

$$c = 1.28\frac{\sqrt{0.2 \cdot 0.8}}{\sqrt{50}} + 0.2 = 0.2724$$

Como podemos ver, a proporção amostral (.27) de usuários que utilizam de forma indevida o sistema não nos leva a rejeitar a hipótese nula, conjecturada pelo adminstrador, a 10% de significância, pois não é superior a 0.2724 (isso é equivalente a verificar que $z_c = \sqrt{50}\frac{0,27 - 0,20}{0,20 * 0,80} = 1,23 < 1,28$. O p-valor nesse caso é dado por $p = P(Z > 1,23) = 1 - P(Z < 1,23) = 0,1093$. Portanto, como $p > \alpha$, nÃ£o se rejeita H_0.

Testes de Hipótese

Exemplo

Para se adequar a padrões internacionais, um engenheiro de sistemas quer assegurar que a produção de processadores esteja sob controle, produzindo processadores com velocidade de processamento de 700 unidades de velocidade de processamento (UVP). Para avaliar essa hipótese, decide-se medir a velocidadede processamento numa amostra aleatória de 10 processadores. Ao final da experiência, obtém-se uma média amostral de 690.2 UVP e uma variância de 2128.9 UVP^2.

Testes de Hipótese

Exemplo

O engenheiro assume que as velocidades de processamento são bem modelados por uma distribuição normal com variância desconhecida. Com base nessas informações, ele deve rejeitar a hipótese de que o processo está sob controle, a 0.05 de significância?

Testes de Hipótese

Quando a variância populacional é desconhecida, devemos usar a informação da variância amostral. Contudo, a variância amostral S^2 também está sujeita a flutuações amostrais (isto é, ela é uma variável aleatória) e a distribuição da estatística do teste não é mais Normal.

Queremos testar a hipótese $\mu = 700$ contra $\mu \neq 700$. A estatística do teste é

$$T = \sqrt{n}\left(\frac{\bar{X} - \mu_0}{S}\right)$$

Testes de Hipótese

Como σ^2 (variância populacional) é desconhecida, sob a hipótese nula a estatística do teste tem distribuição T de Student, com $n-1$ graus de liberdade. Os valores tabelados da distribuição encontram-se disponíveis no site da disciplina
(`https://www.ime.unicamp.br/~cnaber/Tabela%20fda%20T.pdf`).

Note ainda que a hipótese alternativa é bilateral, ou seja, $H_1 : \mu \neq 700$. Então rejeitamos H_0 se $|T| \geq t_{(1-\frac{\alpha}{2}),n-1}$, onde $t_{(1-\frac{\alpha}{2}),n-1}$ é o quantil da distribuição t. Dessa forma, para obter um erro do tipo I de 0.05, admite-se que quando a hipótese nula é verdadeira, podemos errar tanto "para cima" quanto "para baixo", com 0.025 de probabilidade em cada direção.

Testes de Hipótese

Temos que $s^2 = 2128.9$, logo $s = 46.14$. Para $\alpha = 0.05$, $1 - \alpha/2 = 0.975$. Então devemos considerar o quantil 0.975 da distribuição T com 9 graus de liberdade (pois $n = 10$).

Consultando a tabela, obtemos um valor igual a 2.262.

Tabela distribuição t-student

n denota os graus de liberdade
F *denota o valor da distribuição acumulada

Turmas coordenadas ME203, ME180, ME480 e ME414

UNICAMP, 1º semestre 2010

$n \setminus F$	0.75	0.9	0.95	0.975	0.99	0.995	0.9995
1	1.000	3.078	6.314	12.706	31.821	63.657	636.619
2	0.816	1.886	2.920	[illegible]	6.965	9.925	31.599
3	0.765	1.638	2.353	[illegible]	4.541	5.841	12.924
4	0.741	1.533	2.132	[illegible]	3.747	4.604	8.610
5	0.727	1.476	2.015	[illegible]	3.365	4.032	6.869
6	0.718	1.440	1.943	[illegible]	3.143	3.707	5.959
7	0.711	1.415	1.895	[illegible]	2.998	3.499	5.408
8	0.706	1.397	1.860	[illegible]	2.896	3.355	5.041
9	[illegible]	[illegible]	[illegible]	2.262	2.821	3.250	4.781
10	0.700	1.372	1.812	2.228	2.764	3.169	4.587
11	0.697	1.363	1.796	2.201	2.718	3.106	4.437
12	0.695	1.356	1.782	2.179	2.681	3.055	4.318
13	0.694	1.350	1.771	2.160	2.650	3.012	4.221
14	0.692	1.345	1.761	2.145	2.624	2.977	4.140

Testes de Hipótese

Reunindo todas as informações, temos que

$$|T| = \left|\frac{690.2 - 700}{46.14/\sqrt{10}}\right| = |-0.617| < 2.262$$

Então não rejeitamos $H_0 : \mu_A = 700$.
Nesse caso, o p-valor é dado por
$p = 2(1 - P(T < 0,617)|H_0) = 2(1 - 0,75) = 0,5$, em que $T \sim t_{(9)}$. Assim, como $p \geq \alpha$, não rejeitamos H_0 ao nível de significância de 5%.