MBA em Ciência de Dados
Aula 5 – Exemplos
Prof. Francisco Rodrigues

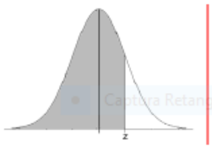
1 – (Morettin, Estatística Básica) Um fabricante de lajotas de cerâmica introduz um novo material em sua fabricação e acredita que aumentará a resistência média, que é de 206 kg. A resistência das lajotas tem distribuição normal com desvio padrão de 12 kg. Retira-se uma amostra de 30 lajotas, obtendo $\bar{x}_{OBS}$ = 210 kg. Ao nível de 10%, pode o fabricante aceitar que a resistência média de suas lajotas tenha aumentado?

RESOLUÇÃO

$$H_0: \mu = 206$$
$$H_a: \mu > 206$$
$$P(\bar{X} > \bar{x}_c | \mu = 206) = 0,10$$
$$P\left(Z > \frac{\bar{x}_c - \mu}{\frac{\sigma}{\sqrt{n}}} \,|\, \mu = 206\right) = 0,10$$
$$P\left(Z > \frac{\bar{x}_c - 206}{\frac{12}{\sqrt{30}}}\right) = 0,10$$

Usando a tabela, temos que P(Z>z) = 0,10 é o mesmo que P(Z < z) = 0,90.
Pela tabela (procurem no corpo da tabela o valor mais próximo de 0,90 e depois verifique qual o valor correspondente de z):

Standard Normal Cumulative Probability Table

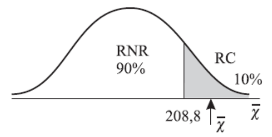


Cumulative probabilities for POSITIVE z-values are shown in the following table:

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	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
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	0.9279	0.9292	0.9306	0.9319

$$\frac{\bar{x}_c - 206}{\frac{12}{\sqrt{30}}} = 1,28$$

Resolvendo: $\bar{x}_c = 214,86$
Como esse valor é maior do que $\bar{x}_{OBS} = 210\ Kg$, rejeitamos H_0 ao nível 10%.



2 - (Morettin, Estatística Básica) Uma fábrica de automóveis anuncia que seus carros consomem, em média, 11 litros por 100 km, com desvio padrão de 0,8 litro. Uma revista decide testar essa afirmação e analisa 35 carros dessa marca, obtendo 11,4 litros por 100 km, como consumo médio. Admitindo que o consumo tenha distribuição normal, ao nível de 10%, o que a revista concluirá sobre o anúncio da fábrica?

RESOLUÇÃO

$$H_0: \mu = 11$$
$$H_a: \mu \neq 11$$
$$P(\bar{X} < \bar{x}_{c1} \cup \bar{X} > \bar{x}_{c2} | \mu = 11) = 0{,}10$$
$$P\left(Z < \frac{\bar{x}_{c1} - 11}{\frac{0{,}8}{\sqrt{35}}}\right) + P\left(Z > \frac{\bar{x}_{c2} - 11}{\frac{0{,}8}{\sqrt{35}}}\right) = 0{,}10$$

Usando a simetria da distribuição normal, temos: $P(Z < z_1) = 0{,}05$ e $P(Z > z_2) = 0{,}95$. Pela tabela, temos: $z_1 = -1{,}64$ e $z_2 = 1{,}64$.

Standard Normal Cumulative Probability Table

Cumulative probabilities for NEGATIVE z-values are shown in the following table:

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
-3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
-3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
-3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
-3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
-2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
-2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
-2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
-2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
-2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
-2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
-2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
-2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
-2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
-1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
-1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
-1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
-1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559

1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767

Resolvendo:

$$\frac{\bar{x}_{c1} - 11}{\frac{0{,}8}{\sqrt{35}}} = -1{,}64$$

$$\frac{\bar{x}_{c2} - 11}{\frac{0,8}{\sqrt{35}}} = 1,64$$

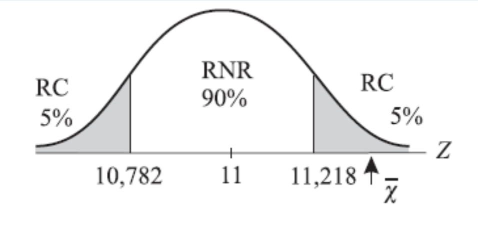
Obtemos:

$$\bar{x}_{c1} = 10,782 \; litros$$
$$\bar{x}_{c2} = 11,218 \; litros$$

Como

$$\bar{x}_{OBS} = 11,4 \; litros,$$

Rejeitamos H_0 ao nível 10%, pois o valor médio observado está na região de rejeição:



3 – (Morettin & Bussab) Os novos operários de uma empresa são treinados a operarem uma máquina, cujo tempo X (em horas) de aprendizado é anotado. Observou-se que X segue de perto a distribuição N(25; 100). Uma nova técnica de ensino, que deve melhorar o tempo de aprendizado, foi testada em 16 novos empregados, os quais apresentaram 20:5 horas como tempo médio de aprendizado. Usando o p-valor, você diria que a nova técnica é melhor que a anterior?

RESOLUÇÃO

Vamos formular as hipóteses:

$$H_0: \mu = 25 \; (tempo\ continua\ igual)$$
$$H_a: \mu < 25$$

A probabilidade de significância (ou p-valor) é obtida calculando-se a probabilidade do valor observado na estatística do teste, ou seja,

$$P(\bar{X} < 20,5 | \mu = 25) = P\left(\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} < \frac{20,5 - 25}{\frac{10}{\sqrt{16}}}\right) = P(Z < -1,8) = 0,036.$$

Neste caso, o p-valor é de apenas 3,6%, o que nos diz que para qualquer nível de significância maior que 3,6%, rejeitamos a hipótese nula). Logo, a nova metodologia parece ser melhor que a anterior.