

Testes de Hipóteses II

O p-valor, e testes com duas amostras

Felipe Figueiredo

Instituto Nacional de Traumatologia e Ortopedia

Sumário

Recapitulando

Vimos como formular hipóteses estatísticas seguindo o procedimento abaixo:

Teste de hipóteses

- 1 Formular as hipóteses nula e alternativa
- 2 Identificar a região crítica (região de rejeição)
- 3 Calcular a estatística de teste adequada
- 4 Rejeitar ou não a hipótese nula

- Este processo sistemático pode ser aplicado a diversos tipos de hipóteses em estudos com dados quantitativos.
- Atualmente tem se usado com mais frequência uma metodologia equivalente usando o **p-valor** (ou valor P).
- Diferença: ao invés de comparar diretamente os Z-escores (região crítica sob a curva), vamos comparar as probabilidades destes (significância)
- Envolve premissas sutis e a interpretação deve ser tomada cuidadosamente (veja artigos complementares no site).

O p-valor



Testes de Hipóteses II

Felipe Figueiredo

Definition

Assumindo que a hipótese nula seja verdadeira, o **p-valor** de um teste de hipóteses é a probabilidade de se obter uma estatística amostral com valores tão extremos, ou mais extremos que aquele observado.

O p-valor **é**:

- Uma estatística (i.e., depende da amostra - dados e tamanho)
- A probabilidade (condicional) de se observar o resultado ao acaso **dado que** a H_0 é verdadeira.
- Uma medida da força da evidência contra a H_0 .

O p-valor



Testes de Hipóteses II

Felipe Figueiredo

Como utilizar

- Quanto menor o p-valor, mais evidências para rejeitar a hipótese nula.
- O ponto de corte mais utilizado é a significância de 5%
- Assim, qualquer $p \leq 0.05$ é estatisticamente significativo.

O p-valor



Testes de Hipóteses II

Felipe Figueiredo

Como calcular

- calcular a estatística de teste apropriada para a situação (teste Z, teste t, etc.)
- encontrar a probabilidade p correspondente a esta estatística (por exemplo, na tabela apropriada, ou com uma ferramenta computacional)
- comparar o p-valor encontrado com a significância do estudo

O p-valor



Testes de Hipóteses II

Felipe Figueiredo



TABLE A: STANDARD NORMAL PROBABILITIES

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
-3.4	.0001	.0003	.0005	.0007	.0009	.0011	.0013	.0015	.0017	.0019
-3.3	.0005	.0007	.0009	.0011	.0013	.0015	.0017	.0019	.0021	.0023
-3.2	.0007	.0009	.0011	.0013	.0015	.0017	.0019	.0021	.0023	.0025
-3.1	.0010	.0012	.0014	.0016	.0018	.0019	.0021	.0023	.0025	.0027
-3.0	.0013	.0015	.0017	.0019	.0021	.0023	.0025	.0027	.0029	.0031
-2.9	.0016	.0018	.0019	.0021	.0023	.0025	.0027	.0029	.0031	.0033
-2.8	.0019	.0021	.0023	.0025	.0027	.0029	.0031	.0033	.0035	.0037
-2.7	.0022	.0024	.0025	.0027	.0029	.0031	.0033	.0035	.0037	.0039
-2.6	.0025	.0027	.0029	.0031	.0033	.0035	.0037	.0039	.0041	.0043
-2.5	.0028	.0030	.0031	.0033	.0035	.0037	.0039	.0041	.0043	.0045
-2.4	.0031	.0033	.0035	.0037	.0039	.0041	.0043	.0045	.0047	.0049
-2.3	.0034	.0036	.0038	.0040	.0042	.0044	.0046	.0048	.0050	.0052
-2.2	.0038	.0040	.0042	.0044	.0046	.0048	.0050	.0052	.0054	.0056
-2.1	.0042	.0044	.0046	.0048	.0050	.0052	.0054	.0056	.0058	.0060
-2.0	.0046	.0048	.0050	.0052	.0054	.0056	.0058	.0060	.0062	.0064
-1.9	.0050	.0052	.0054	.0056	.0058	.0060	.0062	.0064	.0066	.0068
-1.8	.0054	.0056	.0058	.0060	.0062	.0064	.0066	.0068	.0070	.0072
-1.7	.0058	.0060	.0062	.0064	.0066	.0068	.0070	.0072	.0074	.0076
-1.6	.0062	.0064	.0066	.0068	.0070	.0072	.0074	.0076	.0078	.0080
-1.5	.0066	.0068	.0070	.0072	.0074	.0076	.0078	.0080	.0082	.0084
-1.4	.0070	.0072	.0074	.0076	.0078	.0080	.0082	.0084	.0086	.0088
-1.3	.0074	.0076	.0078	.0080	.0082	.0084	.0086	.0088	.0090	.0092
-1.2	.0078	.0080	.0082	.0084	.0086	.0088	.0090	.0092	.0094	.0096
-1.1	.0082	.0084	.0086	.0088	.0090	.0092	.0094	.0096	.0098	.0100
-1.0	.0086	.0088	.0090	.0092	.0094	.0096	.0098	.0100	.0102	.0104
-0.9	.0090	.0092	.0094	.0096	.0098	.0100	.0102	.0104	.0106	.0108
-0.8	.0094	.0096	.0098	.0100	.0102	.0104	.0106	.0108	.0110	.0112
-0.7	.0098	.0100	.0102	.0104	.0106	.0108	.0110	.0112	.0114	.0116
-0.6	.0102	.0104	.0106	.0108	.0110	.0112	.0114	.0116	.0118	.0120
-0.5	.0106	.0108	.0110	.0112	.0114	.0116	.0118	.0120	.0122	.0124
-0.4	.0110	.0112	.0114	.0116	.0118	.0120	.0122	.0124	.0126	.0128
-0.3	.0114	.0116	.0118	.0120	.0122	.0124	.0126	.0128	.0130	.0132
-0.2	.0118	.0120	.0122	.0124	.0126	.0128	.0130	.0132	.0134	.0136
-0.1	.0122	.0124	.0126	.0128	.0130	.0132	.0134	.0136	.0138	.0140
0.0	.0126	.0128	.0130	.0132	.0134	.0136	.0138	.0140	.0142	.0144
0.1	.0130	.0132	.0134	.0136	.0138	.0140	.0142	.0144	.0146	.0148
0.2	.0134	.0136	.0138	.0140	.0142	.0144	.0146	.0148	.0150	.0152
0.3	.0138	.0140	.0142	.0144	.0146	.0148	.0150	.0152	.0154	.0156
0.4	.0142	.0144	.0146	.0148	.0150	.0152	.0154	.0156	.0158	.0160
0.5	.0146	.0148	.0150	.0152	.0154	.0156	.0158	.0160	.0162	.0164
0.6	.0150	.0152	.0154	.0156	.0158	.0160	.0162	.0164	.0166	.0168
0.7	.0154	.0156	.0158	.0160	.0162	.0164	.0166	.0168	.0170	.0172
0.8	.0158	.0160	.0162	.0164	.0166	.0168	.0170	.0172	.0174	.0176
0.9	.0162	.0164	.0166	.0168	.0170	.0172	.0174	.0176	.0178	.0180
1.0	.0166	.0168	.0170	.0172	.0174	.0176	.0178	.0180	.0182	.0184
1.1	.0170	.0172	.0174	.0176	.0178	.0180	.0182	.0184	.0186	.0188
1.2	.0174	.0176	.0178	.0180	.0182	.0184	.0186	.0188	.0190	.0192
1.3	.0178	.0180	.0182	.0184	.0186	.0188	.0190	.0192	.0194	.0196
1.4	.0182	.0184	.0186	.0188	.0190	.0192	.0194	.0196	.0198	.0200
1.5	.0186	.0188	.0190	.0192	.0194	.0196	.0198	.0200	.0202	.0204
1.6	.0190	.0192	.0194	.0196	.0198	.0200	.0202	.0204	.0206	.0208
1.7	.0194	.0196	.0198	.0200	.0202	.0204	.0206	.0208	.0210	.0212
1.8	.0198	.0200	.0202	.0204	.0206	.0208	.0210	.0212	.0214	.0216
1.9	.0202	.0204	.0206	.0208	.0210	.0212	.0214	.0216	.0218	.0220
2.0	.0206	.0208	.0210	.0212	.0214	.0216	.0218	.0220	.0222	.0224
2.1	.0210	.0212	.0214	.0216	.0218	.0220	.0222	.0224	.0226	.0228
2.2	.0214	.0216	.0218	.0220	.0222	.0224	.0226	.0228	.0230	.0232
2.3	.0218	.0220	.0222	.0224	.0226	.0228	.0230	.0232	.0234	.0236
2.4	.0222	.0224	.0226	.0228	.0230	.0232	.0234	.0236	.0238	.0240
2.5	.0226	.0228	.0230	.0232	.0234	.0236	.0238	.0240	.0242	.0244
2.6	.0230	.0232	.0234	.0236	.0238	.0240	.0242	.0244	.0246	.0248
2.7	.0234	.0236	.0238	.0240	.0242	.0244	.0246	.0248	.0250	.0252
2.8	.0238	.0240	.0242	.0244	.0246	.0248	.0250	.0252	.0254	.0256
2.9	.0242	.0244	.0246	.0248	.0250	.0252	.0254	.0256	.0258	.0260
3.0	.0246	.0248	.0250	.0252	.0254	.0256	.0258	.0260	.0262	.0264
3.1	.0250	.0252	.0254	.0256	.0258	.0260	.0262	.0264	.0266	.0268
3.2	.0254	.0256	.0258	.0260	.0262	.0264	.0266	.0268	.0270	.0272
3.3	.0258	.0260	.0262	.0264	.0266	.0268	.0270	.0272	.0274	.0276
3.4	.0262	.0264	.0266	.0268	.0270	.0272	.0274	.0276	.0278	.0280
3.5	.0266	.0268	.0270	.0272	.0274	.0276	.0278	.0280	.0282	.0284
3.6	.0270	.0272	.0274	.0276	.0278	.0280	.0282	.0284	.0286	.0288
3.7	.0274	.0276	.0278	.0280	.0282	.0284	.0286	.0288	.0290	.0292
3.8	.0278	.0280	.0282	.0284	.0286	.0288	.0290	.0292	.0294	.0296
3.9	.0282	.0284	.0286	.0288	.0290	.0292	.0294	.0296	.0298	.0300
4.0	.0286	.0288	.0290	.0292	.0294	.0296	.0298	.0300	.0302	.0304
4.1	.0290	.0292	.0294	.0296	.0298	.0300	.0302	.0304	.0306	.0308
4.2	.0294	.0296	.0298	.0300	.0302	.0304	.0306	.0308	.0310	.0312
4.3	.0298	.0300	.0302	.0304	.0306	.0308	.0310	.0312	.0314	.0316
4.4	.0302	.0304	.0306	.0308	.0310	.0312	.0314	.0316	.0318	.0320
4.5	.0306	.0308	.0310	.0312	.0314	.0316	.0318	.0320	.0322	.0324
4.6	.0310	.0312	.0314	.0316	.0318	.0320	.0322	.0324	.0326	.0328
4.7	.0314	.0316	.0318	.0320	.0322	.0324	.0326	.0328	.0330	.0332
4.8	.0318	.0320	.0322	.0324	.0326	.0328	.0330	.0332	.0334	.0336
4.9	.0322	.0324	.0326	.0328	.0330	.0332	.0334	.0336	.0338	.0340
5.0	.0326	.0328	.0330	.0332	.0334	.0336	.0338	.0340	.0342	.0344
5.1	.0330	.0332	.0334	.0336	.0338	.0340	.0342	.0344	.0346	.0348
5.2	.0334	.0336	.0338	.0340	.0342	.0344	.0346	.0348	.0350	.0352
5.3	.0338	.0340	.0342	.0344	.0346	.0348	.0350	.0352	.0354	.0356
5.4	.0342	.0344	.0346	.0348	.0350	.0352	.0354	.0356	.0358	.0360
5.5	.0346	.0348	.0350	.0352	.0354	.0356	.0358	.0360	.0362	.0364
5.6	.0350	.0352	.0354	.0356	.0358	.0360	.0362	.0364	.0366	.0368
5.7	.0354	.0356	.0358	.0360	.0362	.0364	.0366	.0368	.0370	.0372
5.8	.0358	.0360	.0362	.0364	.0366	.0368	.0370	.0372	.0374	.0376
5.9	.0362	.0364	.0366	.0368	.0370	.0372	.0374	.0376	.0378	.0380
6.0	.0366	.0368	.0370	.0372	.0374	.0376	.0378	.0380	.0382	.0384
6.1	.0370	.0372	.0374	.0376	.0378	.0380	.0382	.0384	.0386	.0388
6.2	.0374	.0376	.0378	.0380	.0382	.0384	.0386	.0388	.0390	.0392
6.3	.0378	.0380	.0382	.0384	.0386	.0388	.0390	.0392	.0394	.0396
6.4	.0382	.0384	.0386	.0388	.0390	.0392	.0394	.0396	.0398	.0400
6.5	.0386	.0388	.0390	.0392	.0394	.0396	.0398	.0400	.0402	.0404
6.6	.0390	.0392	.0394	.0396	.0398	.0400	.0402	.0404	.0406	.0408
6.7	.0394	.0396	.0398	.0400	.0402	.0404	.0406	.0408	.0410	.0412
6.8	.0398	.0400	.0402	.0404	.0406	.0408	.0410	.0412	.0414	.0416
6.9	.0402	.0404	.0406	.0408	.0410	.0412	.0414	.0416	.0418	.0420
7.0	.0406	.0408	.0410	.0412	.0414	.0416	.0418	.0420	.0422	.0424
7.1	.0410	.0412	.0414	.0416	.0418	.0420	.0422	.0424	.0426	.0428
7.2	.0414	.0416	.0418	.0420	.0422	.0424	.0426	.0428	.0430	.0432
7.3	.0418	.0420	.0422	.0424	.0426	.0428	.0430	.0432	.0434	.0436
7.4	.0422	.0424	.0426	.0428	.0430	.0432	.0434	.0436	.0438	.0440
7.5	.0426	.0428	.0430	.0432	.0434	.0436	.0438	.0440	.0442	.0444
7.6	.0430	.0432	.0434	.0436	.0438	.0440	.0442	.0444	.0446	.0448
7.7	.0434	.0436	.0438	.0440	.0442	.0444	.0446	.0448	.0450	.0452
7.8	.0438	.0440	.0442	.0444	.0446	.0448	.0450	.0452	.0454	.0456
7.9	.0442	.0444	.0446	.0448	.0450	.0452	.0454	.0456	.0458	.0460
8.0	.0446	.0448	.0450	.0452	.0454	.0456	.0458	.0460	.0462	.0464
8.1	.0450	.0452	.0454	.0456	.0458	.0460	.0462	.0464	.0466	.0468
8.2	.0454	.0456	.0458	.0460	.0462	.0464	.0466	.0468	.0470	.0472
8.3	.0458	.0460	.0462	.0464	.0466	.0468	.0470	.0472	.0474	.0476
8.4	.0462	.0464	.0466	.0468	.0470	.0472	.0474	.0476	.0478	.0480
8.5	.0466	.0468	.0470	.0472	.0474	.0476	.0478	.0480	.0482	.0484
8.6	.0470	.0472	.0474	.0476	.0478	.0480</				

Exemplo



Testes de
Hipóteses II

Felipe
Figueiredo

Example

Um neurologista está testando o efeito de uma droga no tempo de resposta de um certo estímulo neurológico. Para isto, ele injeta uma dose da droga em **100** ratos, cria os estímulos neurológicos e observa o tempo de resposta em cada animal. O neurologista sabe que o tempo de resposta médio de ratos que não receberam a droga é de **1.2 segundos**. O tempo de resposta médio dos ratos injetados foi de **1.05 segundos**, com desvio padrão amostral de **0.5 segundos**. Você acha que a droga tem efeito no tempo de resposta do estímulo?

Fonte: Khan Academy

Exemplo



Testes de
Hipóteses II

Felipe
Figueiredo

Example

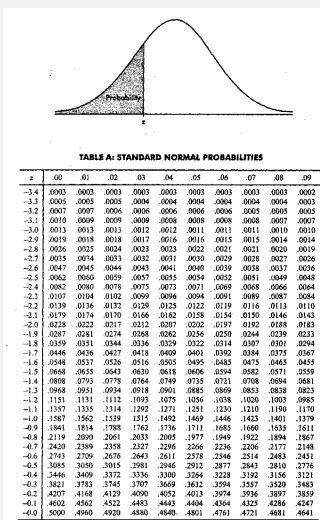
- Dados: $\mu = 1.2$, $\bar{x} = 1.05$, $s = 0.5$, $n = 100$
- $H_0 : \mu = 1.2$, $H_1 : \mu < 1.2$ (teste unicaudal à esquerda)
- n é grande ($n > 30$), então usamos $\sigma \approx s$, e fazemos o teste Z:
- $Z = \frac{1.05 - 1.2}{\frac{0.5}{\sqrt{100}}} = -3$
- Consultando a tabela Z, observamos que este Z-escore corresponde à probabilidade $p = 0.0013$
- Como $p < 0.05$, concluímos que há evidência para rejeitar H_0 .

Tabela Z



Testes de
Hipóteses II

Felipe
Figueiredo



O p-valor



Testes de
Hipóteses II

Felipe
Figueiredo

Cuidado! O p-valor **não é**:

- a probabilidade de que a hipótese nula seja verdadeira
- a probabilidade de que a diferença observada seja devido ao acaso

Estes são erros comuns de interpretação.

O p-valor assume que (1) a hipótese é verdadeira, e (2) que a única causa da diferença é devida ao acaso, portanto não pode ser usado para concluir suas próprias premissas.

"The concept of a p value is not simple and any statements associated with it must be considered cautiously."

Dorey, F. 2010 Clin Orthop Relat Res.

Interpretação do p-valor

- Um valor pequeno para o p-valor (tipicamente $p \leq 0.05$) representa forte evidência para rejeitar a hipótese nula, então deve-se rejeitá-la.
- Um valor alto para o p-valor (tipicamente $p \geq 0.05$) representa pouca evidência contra a hipótese nula, então não se deve rejeitá-la
- Um valor próximo do ponto de corte (0.05) é considerado marginal, portanto “qualquer decisão pode ser tomada”. Sempre apresente seu p-valor para que o leitor possa tirar suas próprias conclusões.

Fonte: Rumsey, D. (Statistics for Dummies, 2nd ed.)

- Frequentemente precisamos dividir os dados em dois grupos e comparar as médias.
- Isto pode ser usado para se estudar o efeito de um tratamento em relação a um grupo controle
- ou mesmo para se comparar dois tratamentos diferentes.

- Para testar a hipótese de que duas médias μ_X e μ_Y são diferentes, consideramos a diferença $\mu_X - \mu_Y$
- Raciocínio: se as médias forem aproximadamente iguais, a diferença será aproximadamente zero
- Procedemos com o teste de hipótese adequado para a situação

Lembre-se que para uma amostra usamos a seguinte estatística de teste:

$$z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

Para duas amostras, é razoável usarmos as estatísticas tanto do grupo 1 quanto do grupo 2.

Estatística de teste:

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sigma_{(\bar{x}_1 - \bar{x}_2)}}$$

onde

$$\sigma_{(\bar{x}_1 - \bar{x}_2)} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

Mas usaremos uma versão simplificada...

Assumindo que H_0 é verdadeira, temos que $\mu_1 = \mu_2 \Rightarrow \mu_1 - \mu_2 = 0$, portanto a estatística de teste que usaremos será:

$$z = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Exemplo

Example

Queremos avaliar a eficiência de uma nova dieta reduzida em gordura no tratamento de obesidade. Selecionamos aleatoriamente 100 pessoas obesas para o grupo 1, que receberão a dieta com pouca gordura. Selecionamos outras 100 pessoas obesas para o grupo 2 que receberão a mesma quantidade de comida, com proporção normal de gordura. Após 4 meses, a perda de peso média no grupo 1 foi de 9.31 lbs ($s=4.67$) e no grupo 2 foi de 7.40 lbs ($s=4.04$). Você acha que essa nova dieta é eficaz na perda de peso?

Fonte: Khan Academy

Exemplo

Example

- Dados: $\bar{x}_1 = 9.31$, $s_1 = 4.67$, $\bar{x}_2 = 7.40$, $s_2 = 4.04$
- $H_0 : \mu_1 = \mu_2 \Rightarrow \mu_1 - \mu_2 = 0 \Rightarrow \mu_{(x_1 - x_2)} = 0$
- $H_1 : \mu_1 - \mu_2 > 0$ (teste unicaudal à direita)
- $\bar{x}_1 - \bar{x}_2 = 1.91$
- $\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{\frac{4.67^2}{100} + \frac{4.04^2}{100}} \approx 0.617$
- Estatística de teste

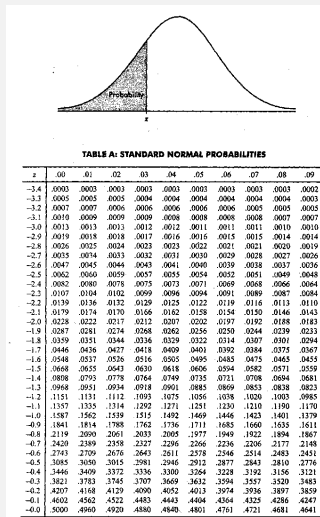
$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{1.91}{0.617} \approx 3.09$$

Tabela Z



Testes de Hipóteses II

Felipe Figueiredo



Exemplo



Testes de Hipóteses II

Felipe Figueiredo

Example

- Encontramos a estatística de teste $z = 3.09$
- Consultando a tabela Z, a probabilidade correspondente é $p = 0.001$
- Como $p < 0.05$, concluímos que há evidências para rejeitar H_0
- Assim, há evidências de que a nova dieta resulta em perda de peso

Amostras pequenas e independentes



Testes de Hipóteses II

Felipe Figueiredo

- Caso uma das amostras seja pequena ($n \leq 30$), deve-se usar o teste t ao invés de teste z.
- A estatística de teste é $t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sigma_{(\bar{x}_1 - \bar{x}_2)}}$
- Cada amostra, n_1 e n_2 tem seus respectivos graus de liberdade
- Consulta-se a tabela t, usando o menor destes valores

Duas amostras pareadas



Testes de Hipóteses II

Felipe Figueiredo

Se duas amostras são pareadas (o **mesmo** indivíduo, antes e depois da intervenção), calculamos a diferença **por indivíduo**, e não a diferença entre as médias de cada grupo.

- Calcula-se a diferença d entre os valores de cada indivíduo
- Considera-se a amostra deste “novo dado”, $H_0 : d = 0$
- A estatística de teste é $t = \frac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n}}}$
- Cada amostra, n_1 e n_2 tem seus respectivos graus de liberdade
- Consulta-se a tabela t, usando o menor destes valores.

Duas amostras grandes e independentes (teste Z):

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sigma_{(\bar{x}_1 - \bar{x}_2)}}$$

Duas amostras pequenas e independentes (teste t):

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sigma_{(\bar{x}_1 - \bar{x}_2)}}$$

Duas amostras dependentes (pareadas):

$$t = \frac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n}}}$$