VE401, Probabilistic Methods in Eng.

Recitation Class - Week 8

Zhanpeng Zhou

UMJI-SJTU Joint Institute

April 20, 2021

Mail: zzp1012@sjtu.edu.cn / WeChat: zzp01012

Table of contents

Sign Test for Median

Sign test. Let $X_1, \ldots, X_n$ be a random sample of size n from an arbitrary continuous distribution and let

$$Q_+ = \#\{X_k : X_k - M_0 > 0\}, \qquad Q_- = \#\{X_k : X_k - M_0 < 0\}.$$

We reject at a significance level α

- $H_0 : M \le M_0$ if $P[Y \le q_- | M = M_0] < \alpha$,
- $H_0 : M \ge M_0$ if $P[Y \le q_+ | M = M_0] < \alpha$,
- $H_0 : M = M_0$ if $P[Y \le \min(q_-, q_+) | M = M_0] < \alpha/2$,

where q_-, q_+ are values of Q_-, Q_+, and Y follows a binomial distri-bution with parameters n' and $1/2$, i.e.,

$$P[Y \le k | M = M_0] = \sum_{y=0}^{k} \binom{n'}{y} \frac{1}{2^{n'}}, \qquad n' = q_+ + q_-.$$

Wilcoxon Signed Rank Test for Median

Wilcoxon signed rank Test. Let $X_1, \ldots, X_n$ be a random sample of size n from a ***symmetric*** distribution. Order the n absolute differences $|X_i - M_0|$ according to the magnitude, so that $X_{R_i} - M_0$ is the R_ith smallest difference by modulus. If ties in the rank occur, the mean of the ranks is assigned to all equal values. Let

$$W_+ = \sum_{R_i > 0} R_i, \qquad |W_-| = \sum_{R_i < 0} |R_i|.$$

We reject at significance level α

- $H_0 : M \leq M_0$ if $|W_-|$ is smaller than the critical value for α,
- $H_0 : M \geq M_0$ if W_+ is smaller than the critical value for α,
- $H_0 : M = M_0$ if $W = \min(W_+, |W_-|)$ is smaller than the critical value for $\alpha/2$.

As is in the sign test, we use n' after discarding data with $X_i = M_0$.

Wilcoxon Signed Rank Test for Median

Normal approximation for distribution of $|W_-|$. Let I_i be the Bernoulli random variable with parameter $1/2$ and $I_i = 1$ if $X_i < M_0$. Then we have

$$
\begin{aligned}
|W_-| = \sum_{i=1}^{n} |R_i| I_i \quad \Rightarrow \quad \mathrm{E}[|W_-|] &= \mathrm{E}\left[\sum_{i=1}^{n} |R_i| I_i\right] \\
&= \sum_{i=1}^{n} \frac{|R_i|}{2} = \frac{n(n+1)}{4}, \\
\operatorname{Var}|W_-| &= \sum_{i=1}^{n} |R_i|^2 \operatorname{Var} I_i \\
&= \sum_{i=1}^{n} \frac{|R_i|^2}{4} = \frac{n(n+1)(2n+1)}{24}.
\end{aligned}
$$

Wilcoxon Signed Rank Test for Median

Normal approximation for distribution of $|W_-|$ (ties). Suppose we have a group of t ties, with ranks R and I given by

$$\{R_{j+1}, \ldots, R_{j+t}\}, \qquad \{I_{j+1}, \ldots, I_{j+t}\}.$$

Suppose for now $R_j > 0$ and denote

$$\overline{R} = \frac{\sum_{k=1}^{t} R_{j+k}}{t} = \frac{2R_{j+1} + t - 1}{2} \;\Rightarrow\; R_{j+1} = \overline{R} - \frac{t-1}{2}.$$

Since the ranks of ties are calculated as the average of the original ranks, the mean does no change. In terms of variance,

$$\begin{aligned} &\sum_{k=1}^{t} |R_{j+k}|^2 \text{Var } I_{j+k} - \sum_{k=1}^{t} |\overline{R}|^2 \text{Var } I_{j+k} \\ =&\frac{1}{4}\left(\sum_{k=1}^{R_{j+1}+t-1} k^2 - \sum_{k=1}^{R_{j+1}-1} k^2 - t\overline{R}^2 \right) =: \frac{1}{4}A. \end{aligned}$$

Wilcoxon Signed Rank Test for Median

Normal approximation for distribution of $|W_-|$ (ties). Then substituting R_{j+1} with $\overline{R}$, we have

$$\begin{aligned} A &= \frac{\left(a+\frac{t}{2}\right)\left(b+\frac{t}{2}\right)(c+t)-\left(a-\frac{t}{2}\right)\left(b-\frac{t}{2}\right)(c-t)}{6} - t\overline{R}^2 \\ &= \frac{t^3-t}{12}. \end{aligned}$$

where

$$a = \overline{R} - \frac{1}{2}, \qquad b = \overline{R} + \frac{1}{2}, \qquad c = 2\overline{R}.$$

Therefore, for each group of t ties, we need to subtract $(t^3 - t)/48$ from the variance. With large sample size, the distribution of $|W_-|$ can be approximated as normal with mean and variance given above.

Wilcoxon Signed Rank Test for Median

Critical values for two-tailed test. For one-tailed test with significance level α, use 2α for lookup.

	alpha values						
n	0.001	0.005	0.01	0.025	0.05	0.10	0.20
5	--	--	--	--	--	0	2
6	--	--	--	--	0	2	3
7	--	--	--	0	2	3	5
8	--	--	0	2	3	5	8
9	--	0	1	3	5	8	10
10	--	1	3	5	8	10	14
11	0	3	5	8	10	13	17
12	1	5	7	10	13	17	21
13	2	7	9	13	17	21	26
14	4	9	12	17	21	25	31
15	6	12	15	20	25	30	36
16	8	15	19	25	29	35	42
17	11	19	23	29	34	41	48
18	14	23	27	34	40	47	55
19	18	27	32	39	46	53	62
20	21	32	37	45	52	60	69
21	25	37	42	51	58	67	77
22	30	42	48	57	65	75	86
23	35	48	54	64	73	83	94
24	40	54	61	72	81	91	104
25	45	60	68	79	89	100	113
26	51	67	75	87	98	110	124
27	57	74	83	96	107	119	134

	alpha values						
n	0.001	0.005	0.01	0.025	0.05	0.10	0.20
28	64	82	91	105	116	130	145
29	71	90	100	114	126	140	157
30	78	98	109	124	137	151	169
31	86	107	118	134	147	163	181
32	94	116	128	144	159	175	194
33	102	126	138	155	170	187	207
34	111	136	148	167	182	200	221
35	120	146	159	178	195	213	235
36	130	157	171	191	208	227	250
37	140	168	182	203	221	241	265
38	150	180	194	216	235	256	281
39	161	192	207	230	249	271	297
40	172	204	220	244	264	286	313
41	183	217	233	258	279	302	330
42	195	230	247	273	294	319	348
43	207	244	261	288	310	336	365
44	220	258	276	303	327	353	384
45	233	272	291	319	343	371	402
46	246	287	307	336	361	389	422
47	260	302	322	353	378	407	441
48	274	318	339	370	396	426	462
49	289	334	355	388	415	446	482
50	304	350	373	406	434	466	503

Estimating Proportions

Proportion. Let $X_1, \ldots, X_n$ be a random sample of X with sample space $\{0, 1\}$, an unbiased estimator for proportion is given by

$$\widehat{p} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

- Statistic and distribution (by central limit theorem).

$$Z = \frac{\widehat{p} - p}{\sqrt{p(1-p)/n}} \sim \text{Normal}(0, 1).$$

- $100(1-\alpha)\%$ two-sided confidence interval for p.

$$\widehat{p} \pm z_{\alpha/2}\sqrt{\widehat{p}(1-\widehat{p})/n}.$$

Estimating Proportions

Proportion. Let $X_1, \ldots, X_n$ be a random sample of X with sample space $\{0, 1\}$, an unbiased estimator for proportion is given by

$$\widehat{p} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

- Choose sample size. $\widehat{p}$ differs from p by at most d with $100(1-\alpha)\%$ confidence.

$$d = z_{\alpha/2}\sqrt{\widehat{p}(1-\widehat{p})/n} \quad \Rightarrow \quad n = \frac{z_{\alpha/2}^2 \widehat{p}(1-\widehat{p})}{d^2}.$$

When no estimate for p is available, we use

$$n = \frac{z_{\alpha/2}^2}{4d^2}.$$

Hypothesis Testing on Proportion

Large-sample test for proportion. Let $X_1, \ldots, X_n$ be a random sample of size n from a Bernoulli distribution with parameter p and let $\widehat{p} = \overline{X}$ denote the sample mean. The test statistic is

$$Z = \frac{\widehat{p} - p_0}{\sqrt{p_0(1 - p_0)/n}}.$$

We reject at significance level α

- $H_0 : p = p_0$ if $|Z| > z_{\alpha/2}$,
- $H_0 : p \le p_0$ if $Z > z_{\alpha}$,
- $H_0 : p \ge p_0$ if $Z < -z_{\alpha}$.

Comparing Two Proportions

Difference of proportions. Suppose we have random samples of sizes n_1, n_2 of $X^{(1)}$ and $X^{(2)}$, respectively.

- Statistic and distribution. For ***large*** sample sizes,

$$Z = \frac{\widehat{p}_1 - \widehat{p}_2 - (p_1 - p_2)}{\sqrt{\dfrac{p_1(1-p_1)}{n_1} + \dfrac{p_2(1-p_2)}{n_2}}} \sim \text{Normal}(0, 1).$$

- $100(1-\alpha)\%$ two-sided confidence interval for $p_1 - p_2$.

$$\widehat{p}_1 - \widehat{p}_2 \pm z_{\alpha/2}\sqrt{\frac{\widehat{p}_1(1-\widehat{p}_2)}{n_1} + \frac{\widehat{p}_2(1-\widehat{p}_2)}{n_2}}.$$

Hypothesis Testing on Difference of Proportions

Test for comparing two proportions. Let $X_1^{(i)}, \ldots, X_{n_i}^{(i)}, i = 1, 2$ be random samples of sizes n_i from two Bernoulli distributions with parameters p_i and let $\widehat{p}_i = \overline{X}_i$ denote the corresponding sample means. The test statistic is given by

$$Z = \frac{\widehat{p}_1 - \widehat{p}_2 - (p_1 - p_2)_0}{\sqrt{\dfrac{\widehat{p}_1(1-\widehat{p}_1)}{n_1} + \dfrac{\widehat{p}_2(1-\widehat{p}_2)}{n_2}}}.$$

We reject at significance level α

- $H_0 : p_1 - p_2 = (p_1 - p_2)_0$ if $|Z| > z_{\alpha/2}$,
- $H_0 : p_1 - p_2 \le (p_1 - p_2)_0$ if $Z > z_\alpha$,
- $H_0 : p_1 - p_2 \ge (p_1 - p_2)_0$ if $Z < -z_\alpha$.

Hypothesis Testing on Equality of Proportions

Pooled test for equality of proportions. Let $X_1^{(i)}, \dots, X_{n_i}^{(i)}, i = 1, 2$ be random samples of sizes n_i from two Bernoulli distributions with parameters p_i and let $\widehat{p}_i = \overline{X}_i$ denote the corresponding sample means. The test statistic is given by

$$Z = \frac{\widehat{p}_1 - \widehat{p}_2}{\sqrt{\widehat{p}(1-\widehat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}, \qquad \widehat{p} = \frac{n_1\widehat{p}_1 + n_2\widehat{p}_2}{n_1 + n_2}.$$

We reject at significance level α

- $H_0 : p_1 = p_2$ if $|Z| > z_{\alpha/2}$,
- $H_0 : p_1 \le p_2$ if $Z > z_\alpha$,
- $H_0 : p_1 \ge p_2$ if $Z < -z_\alpha$.

Basic Distribution

The F-distribution. Let $\chi^2_{\gamma_1}$ and $\chi^2_{\gamma_2}$ be independent chi-squared random variables with γ_1 and γ_2 degrees of freedom, respectively. Then the random variable

$$F_{\gamma_1,\gamma_2} = \frac{\chi^2_{\gamma_1}/\gamma_1}{\chi_{\gamma_2}/\gamma_2}$$

follows a ***F-distribution with γ_1 and γ_2 degrees of freedom***, with density function

$$f_{\gamma_1,\gamma_2} = \gamma_1^{\gamma_1/2}\gamma_2^{\gamma_2/2}\frac{\Gamma\left(\frac{\gamma_1+\gamma_2}{2}\right)}{\Gamma\left(\frac{\gamma_1}{2}\right)\Gamma\left(\frac{\gamma_2}{2}\right)}\frac{x^{\gamma_1/2-1}}{(\gamma_1 x+\gamma_2)^{(\gamma_1+\gamma_2)/2}}$$

for $x \geq 0$ and $f_{\gamma_1,\gamma_2}(x) = 0$ for $x < 0$. Furthermore,

$$P[F_{\gamma_1,\gamma_2} < x] = P\left[\frac{1}{F_{\gamma_1,\gamma_2}} > \frac{1}{x}\right] = 1 - P\left[F_{\gamma_2,\gamma_1} < \frac{1}{x}\right].$$

The F-test for Comparing Variances

F-test. Let S_1^2 and S_2^2 be sample variances based on independent random samples of sizes n_1 and n_2 drawn from normal populations with means μ_1 and μ_2 and variances σ_1^2 and σ_2^2, respectively. The test statistic is given by

$$F_{n_1-1,n_2-1} = \frac{S_1^2}{S_2^2}.$$

We reject at significance level α

- $H_0 : \sigma_1 \le \sigma_2$ if $S_1^2/S_2^2 > f_{\alpha,n_1-1,n_2-1}$,
- $H_0 : \sigma_1 \ge \sigma_2$ if $S_2^2/S_1^2 > f_{\alpha,n_2-1,n_1-1}$,
- $H_0 : \sigma_1 = \sigma_2$ if $S_1^2/S_2^2 > f_{\alpha/2,n_1-1,n_2-1}$ or $S_2^2/S_1^2 > f_{\alpha/2,n_2-1,n_1-1}$.

OC curve. The abscissa is defined by

$$\lambda = \frac{\sigma_1}{\sigma_2}.$$

Comparing Two Means

Basic distribution. Suppose sample means $\overline{X}^{(1)}$ and $\overline{X}^{(2)}$ are cal-culated from samples of sizes n_1 and n_2 respectively from normal populations with means μ_1, μ_2 and variances σ_1, σ_2. Then since

$$\overline{X}^{(1)} \sim \mathrm{N}(\mu_1, \sigma_1^2/n_1), \qquad \overline{X}^{(2)} \sim \mathrm{N}(\mu_2, \sigma_2^2/n_2),$$

the statistic

$$Z = \frac{\overline{X}^{(1)} - \overline{X}^{(2)} - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}$$

follows a standard normal distribution.

Variances Known

Variances known. Let $X_1^{(i)}, \ldots, X_{n_i}^{(i)}$ with $i = 1, 2$ be samples of sizes n_1 and n_2 from normal distributions with unknown means μ_1, μ_2 and ***known*** variances σ_1^2, σ_2^2. Then the test statistic is given by

$$Z = \frac{\overline{X}^{(1)} - \overline{X}^{(2)} - (\mu_1 - \mu_2)_0}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}$$

We reject at significance level α

- $H_0 : \mu_1 - \mu_2 = (\mu_1 - \mu_2)_0$ if $|Z| > z_{\alpha/2}$,
- $H_0 : \mu_1 - \mu_2 \leq (\mu_1 - \mu_2)_0$ if $Z > z_\alpha$,
- $H_0 : \mu_1 - \mu_2 \geq (\mu_1 - \mu_2)_0$ if $Z < -z_\alpha$.

Variances Known

OC curve. When testing equality of means $H_0 : \mu_1 = \mu_2$, we have $(\mu_1 - \mu_2)_0 = 0$. We can use the OC curves for normal distributions with

$$d = \frac{|\mu_1 - \mu_2|}{\sqrt{\sigma_1^2 + \sigma_2^2}}$$

with $n = n_1 = n_2$. When $n_1 \neq n_2$, we use the equivalent sample size

$$n = \frac{\sigma_1^2 + \sigma_2^2}{\sigma_1^2/n_1 + \sigma_2^2/n_2}.$$

Variances Equal but Unknown — Student's T-Test

Variances equal but unknown. Let $X_1^{(i)}, \ldots, X_{n_i}^{(i)}$ with $i = 1, 2$ be samples of sizes n_1 and n_2 from normal distributions with unknown means μ_1, μ_2 and ***equal*** but ***unknown*** variances $\sigma^2 = \sigma_1^2 = \sigma_2^2$. Then the test statistic is given by

$$T_{n_1+n_2-2} = \frac{\overline{X}^{(1)} - \overline{X}^{(2)} - (\mu_1 - \mu_2)_0}{\sqrt{S_p^2(1/n_1 + 1/n_2)}},$$

with ***pooled estimator for variance***

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}.$$

We reject at significance level α

- $H_0 : \mu_1 - \mu_2 = (\mu_1 - \mu_2)_0$ if $|T_{n_1+n_2-2}| > t_{\alpha/2, n_1+n_2-2}$,
- $H_0 : \mu_1 - \mu_2 \leq (\mu_1 - \mu_2)_0$ if $T_{n_1+n_2-2} > t_{\alpha, n_1+n_2-2}$,
- $H_0 : \mu_1 - \mu_2 \geq (\mu_1 - \mu_2)_0$ if $T_{n_1+n_2-2} < -t_{\alpha, n_1+n_2-2}$.

Variances Equal but Unknown — Student's T-Test

OC curve. When testing equality of means $H_0 : \mu_1 = \mu_2$, we have $(\mu_1 - \mu_2)_0 = 0$. We can use the OC curves for the T-test in case of equal sample sizes $n = n_1 = n_2$

$$d = \frac{|\mu_1 - \mu_2|}{2\sigma}.$$

When reading the charts, we must use the modified sample size $n^* = 2n - 1$.

Variances Unequal and Unknown — Welch's T-test

Welch-Satterthwaite Relation. Let $X^{(1)}, \dots, X^{(k)}$ be k independent normally distributed random variables with variances $\sigma_1^2, \dots, \sigma_k^2$. Let $s_1^2, \dots, s_k^2$ be sample variances based on samples of sizes $n_1, \dots, n_k$ from the k populations, respectively. Let $\lambda_1, \dots, \lambda_k > 0$ be positive real numbers and define

$$\gamma := \frac{(\lambda_1 s_1^2 + \cdots + \lambda_k s_k^2)^2}{\displaystyle\sum_{i=1}^{k} \frac{(\lambda_i s_i^2)^2}{n_i - 1}}.$$

Then

$$\chi_\gamma^2 := \gamma \cdot \frac{\lambda_1 s_1^2 + \cdots + \lambda_k s_k^2}{\lambda_1 \sigma_1^2 + \cdots + \lambda_k \sigma_k^2}$$

follows approximately a chi-squared distribution with γ degrees of freedom, where we round γ ***down*** to the nearest integer.

Variances Unequal and Unknown — Welch's T-test

Welch's T-test. Let $X_1^{(i)}, \ldots, X_{n_i}^{(i)}$ with $i = 1, 2$ be samples of sizes n_1 and n_2 from normal distributions with unknown means μ_1, μ_2 and ***unequal*** and ***unknown*** variances σ_1^2, σ_2^2. The test statistic is given by

$$T_\gamma = \frac{\overline{X}^{(1)} - \overline{X}^{(2)} - (\mu_1 - \mu_2)_0}{\sqrt{S_1^2/n_1 + S_2^2/n_2}}, \qquad \gamma = \frac{(S_1^2/n_1 + S_2^2/n_2)^2}{\dfrac{(S_1^2/n_1)^2}{n_1 - 1} + \dfrac{(S_2^2/n_2)^2}{n_2 - 1}}$$

We reject at significance level α

- $H_0 : \mu_1 - \mu_2 = (\mu_1 - \mu_2)_0$ if $T_\gamma > t_{\alpha/2,\gamma}$,
- $H_0 : \mu_1 - \mu_2 \leq (\mu_1 - \mu_2)_0$ if $T_\gamma > t_{\alpha,\gamma}$,
- $H_0 : \mu_1 - \mu_2 \geq (\mu_1 - \mu_2)_0$ if $T_\gamma < -t_{\alpha,\gamma}$.

Wilcoxon Rank-Sum Test

Wilcoxon rank-sum test. Let X and Y be two random populations following some continuous distributions.

Let $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_n$, where $m \leq n$, be random samples from X and Y and associate the rank $R_i, i = 1, \ldots, m+n$, to the R_ith smallest among the $m+n$ total observations. If ties in the rank occur, the mean of the ranks is assigned to all equal values. The test statistic is given by

$$W_m = \text{sum of the ranks of } X_1, \ldots, X_m$$

We reject $H_0 : P[X > Y] = 1/2$ at significance level α if W_m falls into the corresponding critical region.

Wilcoxon Rank-Sum Test

Critical values for Wilcoxon rank-sum test. m is the sample size of the smaller sample, while n is the size of the larger sample. W includes the critical values for two-tailed or one-tailed tests. P is the corresponding p-value.

1-tail 2-tail		$\alpha = 0.025$ $\alpha = 0.05$			$\alpha = 0.05$ $\alpha = 0.10$			1-tail 2-tail		$\alpha = 0.025$ $\alpha = 0.05$			$\alpha = 0.05$ $\alpha = 0.10$		
m	n	W	d	P	W	d	P	m	n	W	d	P	W	d	P
3	3				6 15	1	.0500	5	10	23 57	9	.0200	26 54	12	.0496
3	4				6 18	1	.0286	5	11	24 61	10	.0190	27 58	13	.0449
3	5	6 21	1	.0179	7 20	2	.0357	5	12	26 64	12	.0242	28 62	14	.0409
3	6	7 23	2	.0238	8 22	3	.0476	5	13	27 68	13	.0230	30 65	16	.0473
3	7	7 26	2	.0167	8 25	3	.0333	5	14	28 72	14	.0218	31 69	17	.0435
3	8	8 28	3	.0242	9 27	4	.0424	5	15	29 76	15	.0209	33 72	19	.0491
3	9	8 31	3	.0182	10 29	5	.0500	5	16	30 80	16	.0201	34 76	20	.0455
3	10	9 33	4	.0245	10 32	5	.0385	5	17	32 83	18	.0238	35 80	21	.0425
3	11	9 36	4	.0192	11 34	6	.0440	5	18	33 87	19	.0229	37 83	23	.0472
3	12	10 38	5	.0242	11 37	6	.0352	5	19	34 91	20	.0220	38 87	24	.0442

A larger table can be found in rc files.

Wilcoxon Rank-Sum Test

Wilcoxon rank-sum test. For large values of $m (m \geq 20)$, W_m is approximated normally distributed with

$$\mathrm{E}[W_m] = \frac{m(m+n+1)}{2}, \qquad \mathrm{Var}[W_m] = \frac{mn(m+n+1)}{12}.$$

In case of ties, the variance may be corrected by taking

$$\mathrm{Var}[W_m] = \frac{mn(m+n+1)}{12 - \sum_{\text{groups}} \frac{t^3+t}{12}},$$

where the sum is taken over all groups of t ties.

Variances Equal but Unknown — Paired T-Test

Paired T-test. Let $X_1^{(i)}, \ldots, X_{n_i}^{(i)}$ with $i = 1, 2$ be samples of size $n = n_1 = n_2$ from normal distributions with unknown means μ_1, μ_2 and ***equal*** but ***unknown*** variances $\sigma^2 = \sigma_1^2 = \sigma_2^2$. Then $D_i = X_i - Y_i$ follows normal distributions. Then the test statistic is given by

$$T_{n-1} = \frac{\overline{D} - \mu_0}{\sqrt{S_D^2 / n}}.$$

We reject at significance level α

- $H_0 : \mu_D = \mu_0$ if $|T_{n-1}| > t_{\alpha/2, n-1}$,
- $H_0 : \mu_D \leq \mu_0$ if $T_{n-1} > t_{\alpha, n-1}$,
- $H_0 : \mu_D \geq \mu_0$ if $T_{n-1} < -t_{\alpha, n-1}$.

Paired vs. Pooled T-Tests

With two populations X and Y with equal variances σ^2, we want to test $H_0 : \mu_X = \mu_Y$ using samples of equal size n. Then the statistics are

$$T_{\text{pooled}} = \frac{\overline{X} - \overline{Y}}{\sqrt{2S_p^2/n}}, \qquad \text{critical value} = t_{\alpha/2,2n-2},$$

$$T_{\text{paired}} = \frac{\overline{X} - \overline{Y}}{\sqrt{S_D^2/n}}, \qquad \text{critical value} = t_{\alpha/2,n-1}.$$

Preferring a more powerful test, we consider the following.

- $t_{\alpha/2,2n-2} < t_{\alpha/2,n-1}$, smaller critical values $\Rightarrow$ easier to reject.
- $2S_p^2/n$ estimates $2\sigma^2/n$, while S_D^2/n estimates $\sigma_D^2/n = \sigma_{\overline{D}}^2$, where

$$\sigma_{\overline{D}}^2 = \frac{2\sigma^2}{n}(1 - \rho_{\overline{X}\overline{Y}}) = \frac{2\sigma^2}{n}(1 - \rho_{XY}).$$

When $\rho_{XY} > 0$, paired T-test would be more powerful.

Non-parametric Paired Test

Comparison of medians. Let X and Y be two independent random variables that follow the same distribution but differ only in their location, i.e., $X' := X - \delta$ and Y are independent and identically distributed. Then $D = X - Y$ and $2\delta - D$ follow the same distribution. Therefore, D is symmetric about δ, .i.e.,

$$f_D(\delta + d) = f_D(\delta - d).$$

Then we can perform the Wilcoxon signed-rank test on D.

Estimating Correlation

Estimator for correlation. The unbiased estimators for variance and covariance are given by

$$\widehat{\operatorname{Var}[X]} = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2,$$

$$\widehat{\operatorname{Var}[Y]} = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \overline{Y})^2,$$

$$\widehat{\operatorname{Cov}[X, Y]} = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y}),$$

giving

$$R := \widehat{\rho} = \frac{\sum(X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum(X_i - \overline{X})^2}\sqrt{\sum(Y_i - \overline{Y})^2}}.$$

Hypothesis Tests for the Correlation Coefficient

Distribution. Suppose (X, Y) follows a bivariate normal distribution with relation coefficient $\rho \in (-1, 1)$. For large sample size n, the Fisher transformation of R

$$\frac{1}{2} \ln \left(\frac{1+R}{1-R} \right) = \operatorname{Artanh}(R)$$

is approximately normal with

$$\mu = \frac{1}{2} \ln \left(\frac{1+\rho}{1-\rho} \right) = \operatorname{Artanh}(\rho), \qquad \sigma^2 = \frac{1}{n-3}.$$

Hypothesis Tests for the Correlation Coefficient

Confidence interval. A $100(1-\alpha)\%$ confidence interval for ρ is given by

$$\left[\frac{1+R-(1-R)e^{2z_{\alpha/2}/\sqrt{n-3}}}{1+R+(1-R)e^{2z_{\alpha/2}/\sqrt{n-3}}}, \frac{1+R-(1-R)e^{-2z_{\alpha/2}/\sqrt{n-3}}}{1+R+(1-R)e^{-2z_{\alpha/2}/\sqrt{n-3}}}\right]$$

or

$$\tanh\left(\operatorname{Artanh}(R) \pm \frac{z_{\alpha/2}}{\sqrt{n-3}}\right).$$

Hypothesis Tests for the Correlation Coefficient

Test for correlation coefficient. Suppose $(X_1, Y_1), \ldots, (X_n, Y_n)$ is a sample of size n from a bivariate normal population (X, Y) with correlation coefficient $\rho \in (-1, 1)$. The test statistic is given by

$$\begin{aligned} Z &= \frac{\sqrt{n-3}}{2}\left(\ln\left(\frac{1+R}{1-R}\right) - \ln\left(\frac{1+\rho_0}{1-\rho_0}\right)\right) \\ &= \sqrt{n-3}(\operatorname{Artanh}(R) - \operatorname{Artanh}(\rho_0)). \end{aligned}$$

We reject at significance level α

- $H_0 : \rho = \rho_0$ if $|Z| > z_{\alpha/2}$,
- $H_0 : \rho \le \rho_0$ if $Z > z_\alpha$,
- $H_0 : \rho \ge \rho_0$ if $Z < -z_\alpha$.

1. Test for Statistics
2. Exercises
   - Exercise 1.
   - Exercise 2.
   - Exercise 3.

Exercise 1.

Discuss whether the following interpretations of the P-value are true:

- The P -value represents (an upper bound for) the probability that H_0 is true.
- When H_0 is rejected, the P -value represents (an upper bound for) the probability that H_0 is true even though it was rejected.
- When H_0 is rejected, the P -value represents (an upper bound for) the probability that this rejection occurred even though H_0 was true (Type I error).

Exercise 1.

Discuss whether the following interpretations of the P-value are true:

- The P -value represents (an upper bound for) the probability that H_0 is true.
- When H_0 is rejected, the P -value represents (an upper bound for) the probability that H_0 is true even though it was rejected.
- When H_0 is rejected, the P -value represents (an upper bound for) the probability that this rejection occurred even though H_0 was true (Type I error).

Answers: F, F, T

1. Test for Statistics

2. Exercises
   - Exercise 1.
   - Exercise 2.
   - Exercise 3.

Exercise 2. I

Suppose we have two normally distributed populations $X^{(1)}$ and $X^{(2)}$ with mean μ_1, μ_2 and variances σ_1^2, σ_2^2, respectively. A sample of size $n = 20$ is gathered for each of these populations. We get the sample means and variances as below:

$$\bar{x}_1 = 2.604, \quad \bar{x}_2 = 2.343.$$

$$s_1^2 = 1.263, \quad s_2^2 = 0.534.$$

We want to test the hypotheses

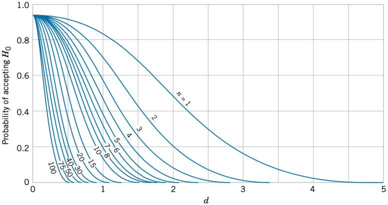
$$H_0 : \mu_1 = \mu_2, \quad H_1 : |\mu_1 - \mu_2| \geq 0.5$$

with significance level $\alpha = 0.05$ in the following cases.

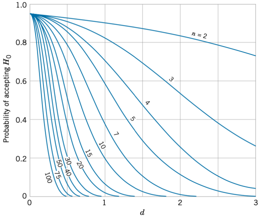
i) We know the variances are $\sigma_1^2 = \sigma_2^2 = 1$. Perform the test. What is the required sample size for the power of the test to be at least 80%?
ii) The variances are unknown but equal $\sigma^2 = \sigma_1^2 = \sigma_2^2$. Perform the test. What is the required sample size for the power of the test to be at least 80%?

Exercise 2. II

iii) The variances are unknown and not necessarily equal. Perform the hypothesis test.



(a) O.C. curves for different values of n for the two-sided normal test for a level of significance $\alpha = 0.05$.



O.C. curves for different values of n for the two-sided t-test for a level of significance $\alpha = 0.05$.

Exercise 2. Answers I

Answers:

1. The test statistic is given by

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} = \frac{2.604 - 2.343}{\sqrt{1/20 + 1/20}} = 0.8254$$

The critical value is given by $z_{\alpha/2} = 1.96 > z$. Therefore, we fail to reject H_0. We calculate

$$d = \frac{|\mu_1 - \mu_2|}{\sqrt{2\sigma^2}} = 0.35$$

and read from OC curve for normal tests. We would require a sample size of at least 75.

Exercise 2. Answers II

ii) We calculate the pooled variance

$$s_p^2 = \frac{(n_1 - 1)\, s_1^2 + (n_2 - 1)\, s_2^2}{n_1 + n_2 - 2} = \frac{s_1^2 + s_2^2}{2} = 0.940.$$

Then the test statistic is given by

$$t_{38} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_p^2\,(1/n_1 + 1/n_2)}} = 0.851$$

The critical value is given by $t_{\alpha/2,38} = 2.024 > t_{38}$. Therefore, we fail to reject H_0. We calculate

$$d = \frac{|\mu_1 - \mu_2|}{2 s_p} = \frac{0.5}{2\sqrt{0.940}} = 0.258$$

where we estimate the variance using pooled variance, and read from OC curve for T -tests. We would require a modified sample size of at least $n^* = 2n - 1 = 75$, giving $n = 38$.

Exercise 2. Answers III

iii. Solution. We calculate

$$\gamma = \frac{\left(s_1^2/n_1 + s_2^2/n_2\right)^2}{\frac{\left(s_1^2/n_1\right)^2}{n_1-1} + \frac{\left(s_2^2/n_2\right)^2}{n_2-1}} = 32.64 \approx 32$$

and thus the test statistic

$$t_{32} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_1^2/n_1 + s_2^2/n_2}} = 0.871$$

with critical value $t_{\alpha/2,32} = 2.04 > t_{32}$. Therefore, we fail to reject H_0

Exercise 3. I

A manufacturer of precision measuring instruments claims that the standard deviation in the use of an instrument is not more than 0.00002 inch. A potential customer requires the instruments to have a standard deviation of not more than 0.00004 inch and will buy them unless there is evidence that the standard deviation is larger than advertised by the manufacturer.

i) Set up appropriate null and alternative hypotheses H_0 and H_1 for a Neyman-Pearson test.

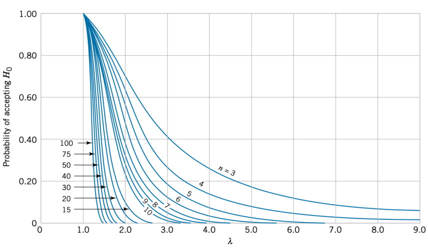
ii) Assuming a normal distribution for the instrument readings. If eight readings are taken, find the critical region for the test, using $\alpha = 0.01$.

iii) What is the power of the test if eight sample readings are taken and H_1 is true?

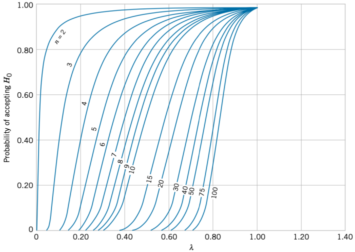
iv) What is the smallest sample size that can be used to detect a true standard deviation of 0.00004 inch or more with a probability of at least 0.95?

Exercise 3. II

v) Eight readings are taken and a sample standard deviation of 0.00005 inch is obtained. What is your conclusion? What is the probability that your decision (to accept either H_0 or H_1) is wrong?



(l) O.C. curves for different values of n for the one-sided (upper tail) chi-square test for a level of significance α = 0.01.



(n) O.C. curves for different values of n for the one-sided (lower tail) chi-square test for a level of significance α = 0.01.

Exercise 3. Answers I

i) We test the hypotheses

$$H_0 : \sigma \leq 0.00002, \quad H_1 : \sigma > 0.00004.$$

ii) If H_0 is true, the statistic

$$X_{n-1}^2 = (n-1)\frac{S^2}{\sigma_0^2}$$

follows a chi-squared distribution with $n-1=7$ degrees of freedom. For $\alpha = 0.01$, the critical value is $\chi^2_{0.01,7} = 18.5$, so the critical region is the interval $(18.5, \infty)$.

Exercise 3. Answers II

iii) We use the OC curve for the right-tailed chi-squared test (why?). The abscissa parameetr is

$$\lambda = \frac{\sigma}{\sigma_0} = \frac{0.00004}{0.00002} = 2$$

and the sample size is $n = 8$. We read off $\beta \approx 0.34$, so the power is approximately $1 - \beta = 0.66$.

iv) Again, we use the OC chart with $\lambda = 2$ and $\beta = 1 - 0.95 = 0.05$. A sample size of $n = 20$ is sufficient to achieve the power stated.

v) The value of the statistic is

$$x_7^2 = 7 \cdot \frac{25 \cdot 10^{-8}}{4 \cdot 10^{-8}} = 43.75$$

Since this lies in the critical region, we reject H_0. There is evidence that the manufacturer's claim is not justified and the customer will not buy instruments. There is less than 1% (P-value) chance that the

Exercise 3. Answers III

decision is not justified and the manufacturer's stated precision is accurate.