VE401 Probabilistic Methods in Eng.
RC 5

CHEN Xiwen

UM-SJTU Joint Institute

April 13, 2020

Table of contents

Fisher's Null Hypothesis Test

Overview.

1. Set up a ***null hypothesis*** H_0 that compares a population parameter θ to a given null value θ_0.
   - $H_0 : \theta = \theta_0$,
   - $H_0 : \theta \leq \theta_0$,
   - $H_0 : \theta \geq \theta_0$.
2. Try to reject the null hypothesis by finding ***P-value*** for the test.
   - One-tailed: upper bound of probability of obtaining the data or more extreme data (based on the null hypothesis), given that the null hypothesis is true.

     $$P[D|H_0] \leq P\text{−value}.$$

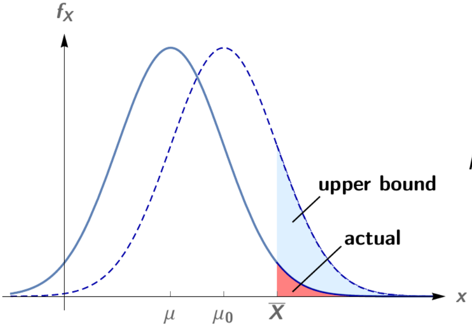
   - Two-tailed: twice of p-value for one-tailed test.
3. We either
   - fail to reject H_0 or
   - reject H_0 at the [p-value] level of significance.

One-tailed Test

Null hypothesis.

$$H_0 : \theta \leq \theta_0 \qquad \text{or} \qquad H_0 : \theta \geq \theta_0.$$

Test for mean. Suppose the sample mean $\overline{X}$ follows a normal distribution with mean μ.



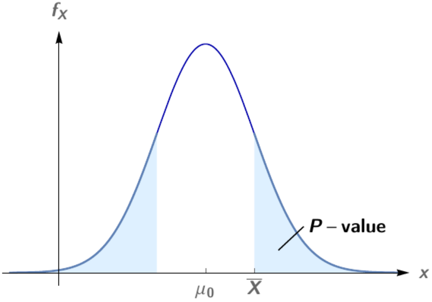
$H_0 : \mu \leq \mu_0$

Two-tailed Test

Null hypothesis.

$$H_0 : \theta = \theta_0.$$

Test for mean. Suppose the sample mean $\overline{X}$ follows a normal distribution with mean μ.



$$H_0 : \mu = \mu_0$$

Neyman-Pearson Decision Theory

Overview.

1. Set up a ***null hypothesis*** H_0 and an ***alternative hypothesis*** H_1.
2. Determine a desirable α and β, where
   - $\alpha := P[\text{reject } H_0 | H_0 \text{ true}]$,
   - $\beta := P[\text{accept } H_0 | H_1 \text{ true}]$, and
   - power $:= 1 - \beta = P[\text{reject } H_0 | H_1 \text{ true}]$.
3. Use α and β to determine the appropriate sample size n. $\triangle$
4. Use α and n to determine the critical region. $\triangle$
5. Obtain sample statistics, and reject H_0 at significance level α and accept H_1 if the test statistic falls into critical region. Otherwise, accept H_0.

Choosing the Sample Size

Normal case. Suppose the sample mean $\overline{X}$ follows a normal distribution with unknown mean μ and known variance σ^2, and we have hypothesis

$$H_0 : \mu = \mu_0, \qquad H_1 : |\mu - \mu_0| \geq \delta_0.$$

Relation between α, β δ, σ and n. With true mean $\mu = \mu_0 + \delta$, the test statistic $Z = \dfrac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} \sim \mathrm{N}(\delta\sqrt{n}/\sigma, 1)$.

$$\begin{aligned}
P[\text{fail to reject } H_0 | \mu = \mu_0 + \delta] &= \frac{1}{\sqrt{2\pi}} \int_{-z_{\alpha/2}}^{z_{\alpha/2}} e^{-(t-\delta\sqrt{n}/\sigma)^2/2} \mathrm{d}t \\
&= \frac{1}{\sqrt{2\pi}} \int_{-z_{\alpha/2}-\delta\sqrt{n}/\sigma}^{z_{\alpha/2}-\delta\sqrt{n}/\sigma} e^{-t^2/2} \mathrm{d}t \\
&\approx \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z_{\alpha/2}-\delta\sqrt{n}/\sigma} e^{-t^2/2} \mathrm{d}t \overset{!}{=} \beta,
\end{aligned}$$

where we set $-z_\beta = z_{\alpha/2} - \delta\sqrt{n}/\sigma$.

Choosing the Sample Size

Normal case. Suppose the sample mean $\overline{X}$ follows a normal distribution with unknown mean μ and known variance σ^2, and we have hypothesis

$$H_0 : \mu = \mu_0, \qquad H_1 : |\mu - \mu_0| \geq \delta_0.$$

Choosing the sample size n.

$$n \approx \frac{(z_{\alpha/2} + z_\beta)^2 \sigma^2}{\delta^2},$$

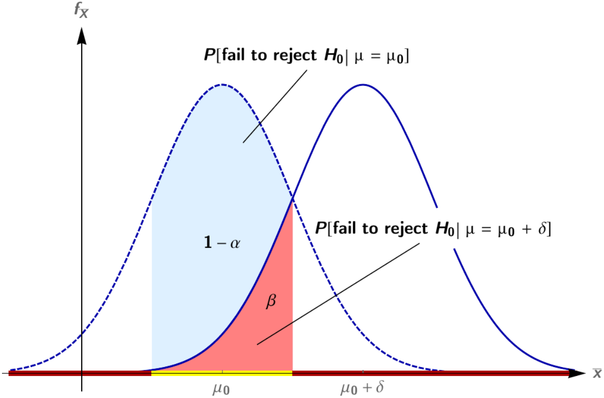
where $z_{\alpha/2}$ and z_β satisfies that

$$\Phi(z_{\alpha/2}) = 1 - \alpha/2, \qquad \Phi(z_\beta) = 1 - \beta,$$

given cumulative distribution function Φ of standard normal distribution.

Choosing the Sample Size

Normal case.



Choosing the Sample Size

More general case: OC curve.

1. For normal test, calculate

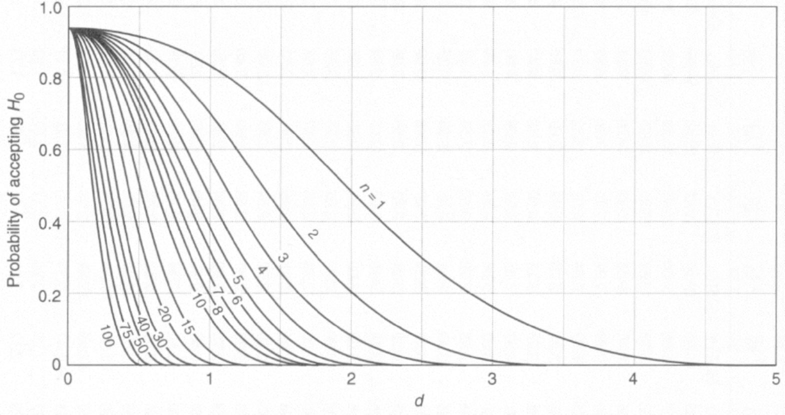
$$d := \frac{|\mu - \mu_0|}{\sigma}.$$

   Note. The abscissa might change corresponding to the distribution of test.

2. Look up in OC curve for sample size n.

Choosing the Sample Size

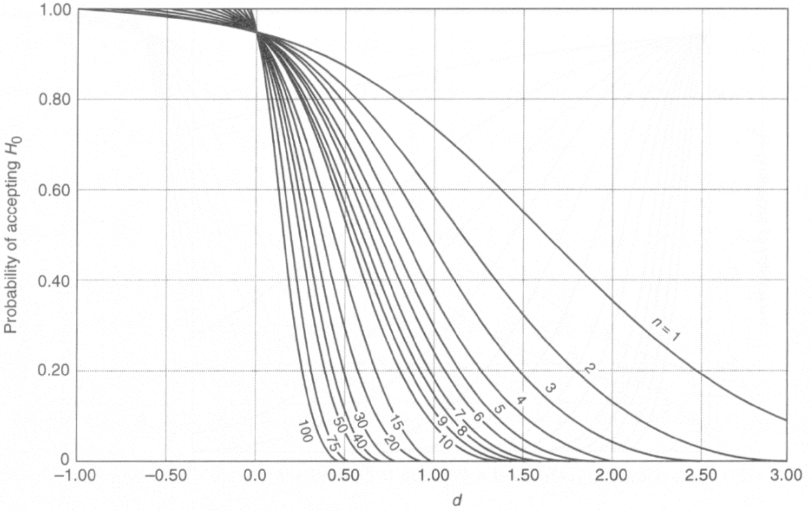
More general case: OC curve.



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

Choosing the Sample Size

More general case: OC curve.



(c) OC curves for different values of n for the one-sided normal test for a level of significance $\alpha = 0.05$.

Choosing the Critical Region

Determine the critical region using α and n. The ***critical region*** is chosen so that if H_0 is true, then the probability of test statistic's value falling into the critical region is no more than α.

Critical region for mean. Suppose the sample mean $\overline{X}$ follows a normal distribution with unknown mean μ and known variance σ^2, with $H_0 : \mu = \mu_0$. Then the test statistic

$$Z = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} \sim \mathrm{N}(0, 1),$$

and thus the critical region is obtained from

$$\frac{|\overline{X} - \mu_0|}{\sigma/\sqrt{n}} > z_{\alpha/2}.$$

We reject H_0 at significance level α and accept H_1 if $\overline{X}$ falls in this critical region.

Neyman-Pearson Decision Theory

Example (assignment 5.3). (... long description ...) ... a researcher would like to test the hypotheses

$$H_0 : \mu \leq 4 \text{ hours}, \qquad H_1 : \mu \geq 4.5 \text{ hours}.$$

A random sample of 50 battery packs is selected and subjected to a life test. Assume that the battery life is normally distributed with standard deviation $\sigma = 0.2$ hours.

Neyman-Pearson Decision Theory

Example (assignment 5.3). (... long description ...) ... a researcher would like to test the hypotheses

$$H_0 : \mu \leq 4 \text{ hours}, \qquad H_1 : \mu \geq 4.5 \text{ hours}.$$

A random sample of 50 battery packs is selected and subjected to a life test. Assume that the battery life is normally distributed with standard deviation $\sigma = 0.2$ hours.

(i) Using the sample mean life span $\overline{X}$ as a test statistic, what is the critical region if $\alpha = 5\%$ is desired.

$$\frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} > z_\alpha \quad \Rightarrow \quad \overline{X} > \mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}} =: L_l.$$

Neyman-Pearson Decision Theory

Example (assignment 5.3). (... long description ...) ... a researcher would like to test the hypotheses

$$H_0 : \mu \leq 4 \text{ hours}, \qquad H_1 : \mu \geq 4.5 \text{ hours}.$$

A random sample of 50 battery packs is selected and subjected to a life test. Assume that the battery life is normally distributed with standard deviation $\sigma = 0.2$ hours.

(ii) Find the power of the test, i.e., the probability of rejecting H_0 if H_1 is true.

$$\begin{aligned}
\text{power} &= P[\overline{X} > L_l | \mu \geq 4.5, \sigma = 0.2] \\
&\geq P[\overline{X} > L_l | \mu = 4.5, \sigma = 0.2] \\
&= P\left[\frac{\overline{X} - 4.5}{\sigma/\sqrt{n}} > \frac{L_l - 4.5}{\sigma/\sqrt{n}}\right].
\end{aligned}$$

Neyman-Pearson Decision Theory

Example (assignment 5.3). (... long description ...) ... a researcher would like to test the hypotheses

$$H_0 : \mu \leq 4 \text{ hours}, \qquad H_1 : \mu \geq 4.5 \text{ hours}.$$

A random sample of 50 battery packs is selected and subjected to a life test. Assume that the battery life is normally distributed with standard deviation $\sigma = 0.2$ hours.

(iii) What sample size would be required to obtain a power of at least 0.97?

$$\alpha = 0.05, \qquad \beta = 1 - 0.97 = 0.03, \qquad \delta = 0.5$$

$$\Rightarrow n \approx \frac{(z_\alpha + z_\beta)^2 \sigma^2}{\delta^2}.$$

Neyman-Pearson Decision Theory

Example (assignment 5.3). (... long description ...) ... a researcher would like to test the hypotheses

$$H_0 : \mu \leq 4 \text{ hours}, \qquad H_1 : \mu \geq 4.5 \text{ hours}.$$

A random sample of 50 battery packs is selected and subjected to a life test. Assume that the battery life is normally distributed with standard deviation $\sigma = 0.2$ hours.

(iv) The sample mean life span turns out to be $\overline{x} = 4.05$ hours. Is H_0 rejected? Find a confidence interval for μ.

Since $\overline{x} > L_l$, H_0 is rejected. A 95% confidence interval for μ is given by

$$\mu \geq \overline{x} - z_\alpha \frac{\sigma}{\sqrt{n}}.$$

Confidence Interval vs. Critical Region

Suppose we would like to estimate the mean μ of a sample $X_1, \ldots, X_n$ of size n.

- Confidence interval. Given a sample data with specific values, the CI gives an interval for the unknown mean μ.
- Critical region. Given a null value μ_0, the critical region gives an interval for sample mean $\overline{X}$ before obtaining specific values.
- Relation. The null hypothesis H_0 is rejected $\Leftrightarrow$ $\overline{X}$ lies in the critical region $\Leftrightarrow$ null value μ_0 lies outside the confidence interval.

Tossing a Coin

Setup. Suppose we have coin, without knowing whether it is fair or not. Let the probability of head be p for the Bernoulli random variable, and we wish to test the hypotheses

$$H_0 : p \leq p_0, \qquad H_1 : p > p_0.$$

Discussion. Following the discussion above, we might toss the coin for n times, and gather a set of sample $X_1, \ldots, X_n$, where each X_i is a Bernoulli random variable. Let $Y = \sum_{i=1}^{n} X_i$. Given a desired significance level α, let $y_\alpha \in \mathbb{N}$ such that

$$Pr[Y \geq y_\alpha | p = p_0] = \sum_{y=y_\alpha}^{n} \binom{n}{y} p_0^y (1-p_0)^{n-y} \leq \alpha,$$

$$Pr[Y \geq y_\alpha - 1 | p = p_0] = \sum_{y=y_\alpha - 1}^{n} \binom{n}{y} p_0^y (1-p_0)^{n-y} > \alpha.$$

Then we would reject H_0 at significance level α if $Y \geq y_\alpha$.

Tossing a Coin

Setup (with prior). Suppose we have coin, without knowing whether it is fair or not. Let the probability of head be p for the Bernoulli random variable, and we wish to test the hypotheses

$$H_0 : p \leq p_0, \qquad H_1 : p > p_0.$$

Now suppose we have an additional prior information. It is learned from the factory which is producing this type of coin that the parameter p follows a Beta distribution with density function

$$f_P(p) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1-p)^{\beta-1}, \qquad p \in (0,1),$$

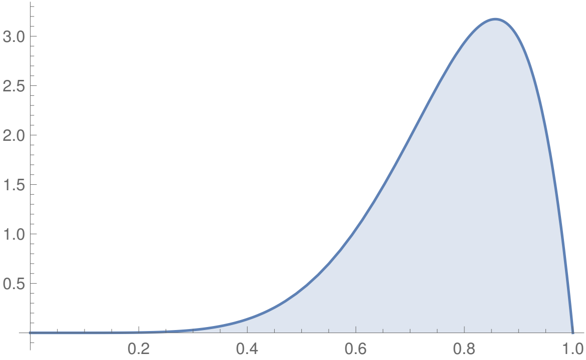
where $\alpha = 7, \beta = 2$. Assume we want to test the hypothesis when the null value $p_0 = 0.5$.

Tossing a Coin

Setup.

In[ ]:=
```
Plot[PDF[BetaDistribution[7, 2], x], {x, 0, 1}, Filling → Axis]
```

Out[ ]=



Tossing a Coin

Bayesian analysis. We have Bayes's theorem (for density function)

$$P[A|B] = \frac{P[B|A] \cdot P[A]}{P[B]} = \frac{P[B, A]}{P[B]}.$$

Then in terms of cumulative distribution function, with $\varepsilon > 0$,

$$\begin{aligned}
F_{Y|X}(y|x) &= \lim_{\varepsilon \to 0} P[Y \leq y | x < X \leq x + \varepsilon] \\
&= \lim_{\varepsilon \to 0} \frac{P[Y \leq y, x < X \leq x + \varepsilon]}{P[x < X \leq x + \varepsilon]} \\
&= \lim_{\varepsilon \to 0} \frac{F_{XY}(x + \varepsilon, y) - F_{XY}(x, y)}{F_X(x + \varepsilon) - F_X(x)} = \frac{1}{f_X(x)} \frac{\partial F_{XY}(x, y)}{\partial x}.
\end{aligned}$$

$$\Rightarrow f_{Y|X}(y|x) = \frac{\partial F_{Y|X}(y|x)}{\partial y} = \frac{f_{XY}(x, y)}{f_X(x)} \propto f_{X|Y}(x|y) f_Y(y).$$

Tossing a Coin

Bayesian analysis. With this additional prior information, according to Bayes's theorem, it is more appropriate to calculate

$$
\begin{aligned}
f_{P|Y}(p|y) &\propto f_{Y|P}(y|p) f_P(p) \\
&\propto \binom{n}{y} p^y (1-p)^{n-y} \cdot \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1-p)^{\beta-1} \\
&\propto p^{y+\alpha-1}(1-p)^{n-y+\beta-1}.
\end{aligned}
$$

and thus

$$
f_{P|Y}(p|y) = \frac{\Gamma(\alpha'+\beta')}{\Gamma(\alpha')\Gamma(\beta')} p^{\alpha'-1}(1-p)^{\beta'-1},
$$

where $\alpha' = y + \alpha, \beta' = n - y + \beta$.

Tossing a Coin

Bayesian analysis. Then we are able to calculate $P[H_0|D]$, which we are really interested in, as

$$P[H_0|D] = F_{P|Y}(p_0|y) = \int_0^{p_0} f_{P|Y}(p|y)\mathrm{d}p.$$

Let y_α satisfies that

$$1 - F_{P|Y}(p|y_\alpha) = 1 - \int_0^{p_0} f_{P|Y}(p|y_\alpha)\mathrm{d}p \le \alpha,$$
$$1 - F_{P|Y}(p|y_\alpha - 1) = 1 - \int_0^{p_0} f_{P|Y}(p|y_\alpha)\mathrm{d}p > \alpha.$$

Then finally we would reject $H_0 : p \le p_0$ if $Y \ge y_\alpha$.

Test for Mean (Variance Known)

Z-test. Let $X_1, \ldots, X_n$ be a random sample of size n from a normal distribution with ***unknown*** mean μ and ***known*** variance σ^2. Let μ_0 be a null value of the mean. Then the test statistic is given by

$$Z = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}.$$

We reject at significance level α

- $H_0 : \mu = \mu_0$ if $|Z| > z_{\alpha/2}$,
- $H_0 : \mu \leq \mu_0$ if $Z > z_\alpha$,
- $H_0 : \mu \geq \mu_0$ if $Z < -z_\alpha$.

OC curve. The abscissa is defined by

$$d = \frac{|\mu - \mu_0|}{\sigma}.$$

Test for Mean (Variance Unknown)

T-test. Let $X_1, \ldots, X_n$ be a random sample of size n from a normal distribution with ***unknown*** mean μ and ***unknown*** variance σ^2. Let μ_0 be a null value of the mean. Then the test statistic is given by

$$T_{n-1} = \frac{\overline{X} - \mu_0}{S/\sqrt{n}}.$$

We reject at significance level α

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

OC curve. The abscissa is defined by

$$d = \frac{|\mu - \mu_0|}{\sigma},$$

where in practice, the unknown σ can be substituted by S.

Test for Variance

Chi-squared test. Let $X_1, \dots, X_n$ be a random sample of size n from a normal distribution with unknown variance σ^2. Let σ_0^2 be a null value of the variance. Then the test statistic is given by

$$\chi^2_{n-1} = \frac{(n-1)S^2}{\sigma_0^2}.$$

We reject at significance level α

- $H_0 : \sigma = \sigma_0$ if $\chi^2_{n-1} \in (0, \chi^2_{1-\alpha/2,n-1}) \cup (\chi^2_{\alpha/2,n-1}, \infty)$,
- $H_0 : \sigma \leq \sigma_0$ if $\chi^2_{n-1} > \chi^2_{\alpha,n-1}$,
- $H_0 : \sigma \geq \sigma_0$ if $\chi^2_{n-1} < \chi^2_{1-\alpha,n-1}$.

OC curve. The abscissa is defined by

$$\lambda = \frac{\sigma}{\sigma_0}.$$

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 distribution 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

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

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

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

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,

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

Wilcoxon Signed Rank Test for Median

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

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

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 $X_1, \dots, X_n$ 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 $|Z| > z_{\alpha/2}$,
- $H_0 : p \leq p_0$ if $Z > z_\alpha$,
- $H_0 : p \geq p_0$ 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 $|Z| > z_{\alpha/2}$,
- $H_0 : p_1 - p_2 \leq (p_1 - p_2)_0$ if $Z > z_{\alpha}$,
- $H_0 : p_1 - p_2 \geq (p_1 - p_2)_0$ if $Z < -z_{\alpha}$.

Hypothesis Testing on Equality of Proportions

Pooled test for equality of proportions. Let $X_1^{(i)}, \ldots, X_{n_i}^{(i)}, i = 1, 2$ 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

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

We reject at significance level α

- $H_0 : p_1 = p_2$ if $|Z| > z_{\alpha/2}$,
- $H_0 : p_1 \leq p_2$ if $Z > z_\alpha$,
- $H_0 : p_1 \geq p_2$ 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

- $H_0 : \sigma_1 \leq \sigma_2$ if $S_1^2/S_2^2 > f_{\alpha,n_1-1,n_2-1}$,
- $H_0 : \sigma_1 \geq \sigma_2$ 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

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

Thanks for your attention!