Lecture 14 12/29/2021

主題：雙樣本檢定相關議題

- 統計在生活中應用最廣的部分是用來做“比較”.
  - 哪個藥的療效佳
    - 治療組 vs. 對照組
  - 性別差異是否存在
    - 男 vs. 女
  - 比較城鄉差距
- 這個單元的主題目的對象是比較兩個測度的“平均”
  - Paired data: 一次取兩個成對的測度 (“paired observations”)

    → take the difference (取差值) → 視為單樣本
  - Two samples → 期末考考到這裡
- 補充但不考的部分
  - Checking the normal assumption

    檢驗分配是否來自常態的母體 (t 檢定的基本假設)

    → Use normal plot
  - 當常態假設違反時 → 用無母數檢定法 (視情況補充)
  - Multiple comparison: Need to be careful!!

    由兩組比較 → 多組比較

Review: Two Sample Comparison

背景：兩個常態母體，欲檢驗 $H_0: \mu_1 - \mu_2 = 0$；分別獨立取 n_1 和 n_2 兩個樣本

population 1: normal with mean μ_1 and variance σ_1^2

→ sample 1: $X_{11},\ldots,X_{1n_1}$ (n_1 observations) → find $\bar{X}_1$ and S_1

population 2: normal with mean μ_2 and variance σ_2^2

→ sample 2: $X_{21},\ldots,X_{2n_2}$ (n_2 observations) → find $\bar{X}_2$ and S_2

Additional assumption: the measurements from the two samples are independent

Case 1: $H_0: \mu_1 - \mu_2 = 0$ versus $H_a: \mu_1 - \mu_2 > 0$

Case 2: $H_0: \mu_1 - \mu_2 = 0$ versus $H_a: \mu_1 - \mu_2 < 0$

Case 3: $H_0: \mu_1 - \mu_2 = 0$ versus $H_a: \mu_1 - \mu_2 \neq 0$

Construction of the test statistic:

1. Use $\bar{X}_1 - \bar{X}_2$ as the point estimator of $\mu_1 - \mu_2$
2. 分配理論: when the two populations are normal with mean and variance equal to

   (μ_1, σ_1^2) and (μ_2, σ_2^2) respectively

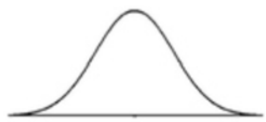
$$\bar{X}_1 - \bar{X}_2 \sim N(\mu_1 - \mu_2, Var(\bar{X}_1 - \bar{X}_2) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_1}) \quad \Leftrightarrow \frac{(\bar{X}_1 - \bar{X}_2) - 0}{\sqrt{\sigma_1^2 / n_1 + \sigma_2^2 / n_2}} \sim N(0,1)$$

- **Two-sample Z test – when σ_1^2 and σ_2^2 are known 幫助理解概念,但不實用**

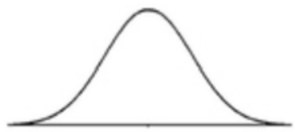
Case 1: $H_0: \mu_1 = \mu_2$ versus $H_a: \mu_1 - \mu_2 > 0$

Method 1: Reject H_0 if $obs\{\bar{X}_1 - \bar{X}_2\} > 0 + z_\alpha \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$

$\Leftrightarrow$ Reject H_0 if $z_{obs} = \frac{obs\{\bar{X}_1 - \bar{X}_2\} - 0}{\sqrt{\sigma_1^2 / n_1 + \sigma_2^2 / n_2}} > z_\alpha$



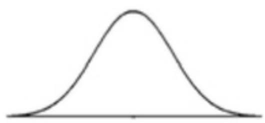
Method 2: p-value = $\Pr(Z > \frac{obs\{\bar{X}_1 - \bar{X}_2\} - 0}{\sqrt{\sigma_1^2 / n_1 + \sigma_2^2 / n_2}}) = \Pr(Z > z_{obs})$



➔ Reject H_0 if p-value < α

Case 2: $H_0: \mu_1 = \mu_2$ versus $H_a: \mu_1 - \mu_2 < 0$

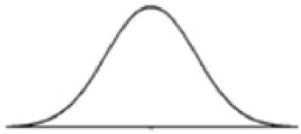
Method 1: reject H_0 if $z_{obs} = \frac{obs\{\bar{X}_1 - \bar{X}_2\} - 0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} < -z_{\alpha/2}$,



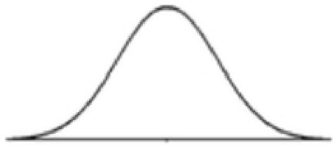
Method 2: p-value = $\Pr(Z < \frac{obs\{\bar{X}_1 - \bar{X}_2\} - 0}{\sqrt{\sigma_1^2 / n_1 + \sigma_2^2 / n_2}})$

Reject if p-value < α

Case 3: $H_0 : \mu_1 = \mu_2$ versus $H_a : \mu_1 - \mu_2 \neq 0$



Method 1: Reject H_0 if $|z_{obs}| = \left| \dfrac{obs\{\bar{X}_1 - \bar{X}_2\} - 0}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \right| > z_{\alpha/2}$



Method 2: p-value $= 2\Pr(Z > |z_{obs}|)$

Reject if p-value < α

Case 3 (method 3): $(1-\alpha)\cdot 100\%$ 信賴區間 for $\mu_1 - \mu_2$:

$$obs\{\bar{X}_1 - \bar{X}_2\} \pm z_{\alpha/2} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\ .$$

Rule: Reject H_0 if 0 falls <u>outside</u> $obs\{\bar{X}_1 - \bar{X}_2\} \pm z_{\alpha/2} \cdot \sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}$

Topic: Two-sample T test – when σ_1^2 and σ_2^2 are both unknown →必考

Fact: When the two populations are normal

$$\frac{\bar{X}_1 - \bar{X}_2}{\sqrt{s_1^2/n_1 + s_2^2/n_2}} \sim^{approximately} T_{v^*},$$

where $v^* = \dfrac{\left(S_1^2/n_1 + S_2^2/n_2\right)^2}{\dfrac{\left(S_1^2/n_1\right)^2}{n_1 - 1} + \dfrac{\left(S_2^2/n_2\right)^2}{n_2 - 1}}$ (not an integer.)

註:

- 即使兩個母體來自常態分佈，只要母體變異數 (σ_1^2, σ_2^2) 是估計的，

  檢定統計量 $\dfrac{\bar{X}_1 - \bar{X}_2}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}$ 也不是真正的 T 分配，而是近似的

  (還伴隨複雜非整數的自由度). → 統計領域一個重要的“常識”. 記住它!

以下介紹克服此問題的幾個選項.

選項一：如果有電腦程式，就可以計算 T_{v^*} 分配並求出 p-value.

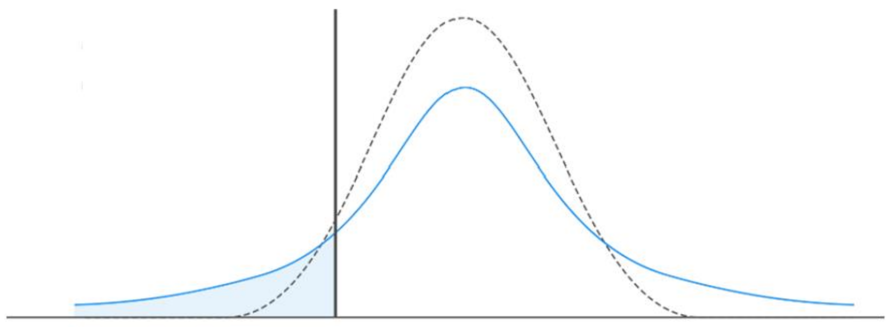
選項二：如果沒有電腦程式，可採取<u>保守原則</u>

Conservative principle: choose $df = \min(n_1 - 1, n_2 - 1)$ →選“樣本少”的自由度

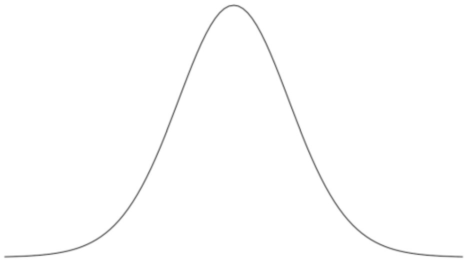
結果：保守原則會使得 $df < v^*$，較不容易拒絕 $H_0: \mu_1 = \mu_2$

保守原則會訂出更嚴格的拒絕條件 (更不易拒絕)

如果在保守原則下，都能拒絕 H_0，代表用真正的自由度，也會拒絕



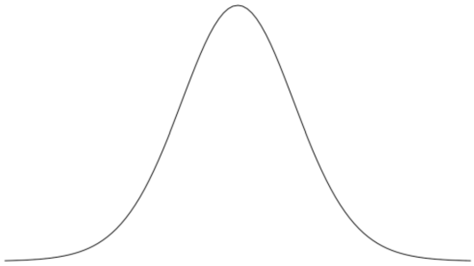
Let $t_{obs} = \dfrac{obs\{\bar{X}_1 - \bar{X}_2\} - 0}{\sqrt{s_1^2 / n_1 + s_2^2 / n_2}}$



$t_{df,\alpha}$ satisfies $\Pr(T_{df} > t_{df,\alpha}) = \alpha$

Case 1: $H_0 : \mu_1 - \mu_2 = 0$ versus $H_a : \mu_1 - \mu_2 > 0$

Method 1: Reject H_0 if $t_{obs} > t_{df,\alpha}$



Method 2: p-value = $\Pr(T_{df} > t_{df,\alpha})$

Reject if p-value < α

Case 2: $H_0 : \mu_1 = \mu_2$ versus $H_a : \mu_1 - \mu_2 < 0$

Method 1: Reject H_0 if $t_{obs} < -t_{df,\alpha}$,

Method 2: p-value = $\Pr(T_{df} < t_{df,\alpha})$

Reject if p-value < α

Case 3: $H_0 : \mu_1 = \mu_2$ versus $H_a : \mu_1 - \mu_2 \neq 0$

Method 1: reject H_0 if $|t_{obs}| > t_{df,\alpha/2}$

Method 2: p-value $= 2\Pr(T_{df} > |t_{obs}|)$

Reject if p-value < α

Method 3: $(1-\alpha)\cdot 100\%$ 信賴區間 for $\mu_1 - \mu_2$:

$$obs\{\bar{X}_1 - \bar{X}_2\} \pm t_{df,\alpha/2} \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}.$$

Reject if 0 falls <u>outside</u> $obs\{\bar{X}_1 - \bar{X}_2\} \pm t_{df,\alpha/2} \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$

EXAMPLE 18.4 Brain size and autism

STATE: Is autism marked by different brain growth patterns in early life, even b the diagnosis is made? Studies have linked brain size in infants and toddlers to a nu of future ailments, including autism. One study looked at the brain sizes of 30 au boys and 12 nonautistic boys (control) who all had received an MRI scan as tod Here are their whole-brain volumes in milliliters:[3]

Autistic									
1311	1250	1292	1419	1401	1297	1202	1336	1308	1353
1515	1461	1365	1364	1362	1303	1278	1247	1333	1340
1319	1286	1223	1241	1229	1209	1171	1154	1128	1230
Control									
1040	1180	1207	1179	1115	1133	1298	1263	1194	1198
1230	1114								

FORMULATE: We had no specific direction for the difference in brain volumes b looking at the data, so the alternative is two-sided. We will test the hypotheses

$$H_0: \mu_1 = \mu_2 \text{ (that is, } \mu_1 - \mu_2 = 0)$$
$$H_a: \mu_1 \neq \mu_2 \text{ (that is, } \mu_1 - \mu_2 \neq 0)$$

Group	Condition	n	$\bar{x}$	s
1	Autistic	30	1297.6	88.4
2	Control	12	1179.3	70.7

The two-sample t statistic is

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} = \frac{1297.6 - 1179.3}{\sqrt{\dfrac{88.4^2}{30} + \dfrac{70.7^2}{12}}} = \frac{118.3}{26.02} = 4.55$$

Degree of freedom $= \min(30-1, 12-1) = 11$, $4.55 > 4.437$ $\Pr(T_{11} > 4.55) < 0.0005$

$p\text{-}value / 2 < 0.0005$ → p-value < 0.001, then Reject H_0

ν	.40 .60	.25 .75	.10 .90	.05 .95	.025 .975	.01 .99	.005 .995	.001 .999	.0005 (Area to right) .9995 (Area to left)	
1	0.325	1.000	3.078	6.314	12.706	31.821	63.657	318.317	636.607	
2	0.289	0.816	1.886	2.920	4.303	6.965	9.925	22.327	31.598	
3	0.277	0.765	1.638	2.353	3.182	4.541	5.841	10.215	12.924	
4	0.271	0.741	1.533	2.132	2.776	3.747	4.604	7.173	8.610	
5	0.267	0.727	1.476	2.015	2.571	3.365	4.032	5.893	6.869	
6	0.265	0.718	1.440	1.943	2.447	3.143	3.707	5.208	5.959	
7	0.263	0.711	1.415	1.895	2.365	2.998	3.499	4.785	5.408	
8	0.262	0.706	1.397	1.860	2.306	2.896	3.355	4.501	5.041	
9	0.261	0.703	1.383	1.833	2.262	2.821	3.250	4.297	4.781	
10	0.260	0.700	1.372	1.812	2.228	2.764	3.169	4.144	4.587	
11	0.260	0.697	1.363	1.796	2.201	2.718	3.106	4.025	4.437	4.55
12	0.259	0.695	1.356	1.782	2.179	2.681	3.055	3.930	4.318	
13	0.259	0.694	1.350	1.771	2.160	2.650	3.012	3.852	4.221	

Example:

紫錐花的3種功效及副作用（6點使用禁忌要小心）- 營養新知
https://formulawave.com/echinacea-benefits/ ▾
2018年7月4日 - 紫錐花（Echinacea）是什麼？ 紫錐花被推薦的實證功效(好處)有哪些？ 1・普通感冒（common cold）; 2・降低復發性呼吸道感染及併發症; 3・
紫錐花（Echinacea）是什麼？・紫錐花被推薦的實證功效 ...・紫錐花有副作用嗎？

Echinacea for the common cold? Echinacea is widely used as an herbal remedy for the common cold, but does it work? In a double-blind experiment, healthy volunteers agreed to be exposed to common-cold-causing rhinovirus type 39 and have their symptoms monitored. The volunteers were randomly assigned to take either a placebo or an echinacea supplement daily from 7 days before till 5 days after viral exposure. A symptom score was recorded for each subject over the 5 days following exposure, with higher scores indicating more severe symptoms. The published results reported the mean ± SEM for both groups as 13.21 ± 1.91 (echinacea) and 15.05 ± 1.43 (placebo).[6]

$\bar{x}_1$

$n_1 = 52, n_2 = 103$ → df = 51

Based on confidence intervals, find the standard deviations.

$$\bar{x}_1 \pm t_{51,0.025}\frac{s_1}{\sqrt{52}} = 13.21 \pm 1.91 \;\rightarrow\; t_{51,0.025} = 2.008 \rightarrow s_1 = \frac{1.91\times\sqrt{52}}{2.008} = 6.859$$

$$\bar{x}_2 \pm t_{102,0.025}\frac{s_2}{\sqrt{102}} = 15.05 \pm 1.43 \;\rightarrow\; t_{102,0.025} \approx 1.96 \rightarrow s_2 = \frac{1.43\times\sqrt{103}}{1.96} = 7.369$$

The compute the observed T statistics:

$$Var(\bar{X}_1 - \bar{X}_2) = \frac{S_1^2}{n_1} + \frac{S_2^2}{n_2} = 1.4319$$

$$t_{obs} = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = -1.56$$ → p-value = $\Pr(T_{51} < -1.56) \in (0.05, 0.1)$

→ fail to reject $H_0: \mu_1 - \mu_2 = 0$ vs. $H_a: \mu_1 - \mu_2 < 0$

選項三：假設兩個樣本獨立，$\sigma_1 = \sigma_2$ 就可以推導出 exact T 分配 (不考)

Two sample study (when $\sigma_1^2 = \sigma_2^2 = \sigma^2$ which is unknown) → 不考

population 1: mean μ_1 and variance σ^2 → 求取平均 $\bar{X}_1$

population 2: mean μ_2 and variance σ^2 →求取平均 $\bar{X}_2$

估計 σ^2 : pool the data → $X_{11},\ldots,X_{1n_1}$, $X_{21},\ldots,X_{2n_2}$ 求合併樣本的變異數

$$s_p^2 = \frac{\sum_{i=1}^{n_1}(X_{1i}-\bar{X}_1)^2 + \sum_{j=1}^{n_2}(X_{2j}-\bar{X}_2)^2}{n_1+n_2-2} \rightarrow \sigma^2 \text{ 的估計量}$$

$$Var(\bar{X}_1-\bar{X}_2) = \sqrt{\frac{\sigma^2}{n_1}+\frac{\sigma^2}{n_2}} = \sigma\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \rightarrow \text{估計量 } s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

Fact: When the two populations are both normal, under $H_0: \mu_1-\mu_2=0$

$$\frac{(\bar{X}_1-\bar{X}_2)-0}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim \text{exactly } T_{n_1+n_2-2} \text{ (自由度很大，所以比較容易拒絕)}$$

Example: Comparing Packing Machines

In a packing plant, a machine packs cartons with jars. It is supposed that a new machine will pack faster on the average than the machine currently used. To test that hypothesis, the times it takes each machine to pack ten cartons are recorded. The results in seconds, are shown in the tables.

New Machine					Old Machine				
42.1	41.3	42.4	43.2	41.8	42.7	43.8	42.5	43.1	44.0
41.0	41.8	42.8	42.3	42.7	43.6	43.3	43.5	41.7	44.1

$\bar{x}_1 = 42.14, s_1 = 0.683$ $\quad\quad$ $\bar{x}_2 = 43.23, s_2 = 0.750$

$$\begin{aligned} s_p &= \sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}} \\ &= \sqrt{\frac{(10-1)(0.683)^2+(10-1)(0.750)^2}{10+10-2}} \\ &= \sqrt{\frac{9.261}{18}} \\ &= 0.7173 \end{aligned}$$

$$\begin{aligned} t^* &= \frac{\bar{x}_1-\bar{x}_2-0}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \\ &= \frac{42.14-43.23}{0.7173\sqrt{\frac{1}{10}+\frac{1}{10}}} \\ &= -3.398 \end{aligned}$$

The alternative is left-tailed so the critical value is the value a such that $P(T<a)=0.05$, with $10+10-2=18$ degrees of freedom. The critical value is -1.7341. The rejection region is $t^*<-1.7341$.

Remark: 假設 $\sigma_1^2=\sigma_2^2=\sigma^2$，去檢定 $H_0:\mu_1=\mu_2$，其實並不合理!

Robustness 討論

萬一常態假設不吻合, T test 或是 T confidence interval 還可以用嗎?

USING THE t PROCEDURES

- Except in the case of small samples, the condition that the data are an SRS from the population of interest is more important than the condition that the population distribution is Normal.
- *Sample size less than 15:* Use t procedures if the data appear close to Normal (roughly symmetric, single peak, no outliers). If the data are skewed or if outliers are present, do not use t.
- *Sample size at least 15:* The t procedures can be used except in the presence of outliers or strong skewness.
- *Large samples:* The t procedures can be used even for clearly skewed distributions when the sample is large, roughly $n \geq 40$.

Robustness again

The two-sample t procedures are more robust than the one-sample t methods, particularly when the distributions are not symmetric. When the sizes of the two samples are equal and the two populations being compared have distributions with similar shapes, probability values from the t table are quite accurate for a broad range of distributions when the sample sizes are as small as $n_1 = n_2 = 5$.[8] When the two population distributions have different shapes, larger samples are needed.

As a guide to practice, adapt the guidelines given on page 446 for the use of one-sample t procedures to two-sample procedures by replacing "sample size" with "sum of the sample sizes," n_1+n_2. These guidelines err on the side of safety, especially when the two samples are of equal size. *In planning a two-sample study, choose equal sample sizes whenever possible. The two-sample t procedures are most robust against non-Normality in this case.*

討論: normality 假設對 t 檢定的影響

- 檢驗 normality assumption 的方法:

a. Normal plot → 一種 quantile-quantile (Q-Q) plot → 期待看到“直線”

b. Stem-leaf plot (或是 histogram) → 看到對稱

c. 尚有正式的 goodness-of-fit 檢定 → 如 K-S test (軟體會讓你選)

補充: 檢驗 normality assumption

主題: normal probability plot

Assumption: $\varepsilon_i \stackrel{iid}{\sim} N(0,\sigma_e^2)$ → MLE needs this assumption

Ideal plot: a straight line

Example: Students'scores behave like a normal distribution?

67 72 74 80 81 83 86 87 90 93

Question: 學生的成績服從常態分配嗎?

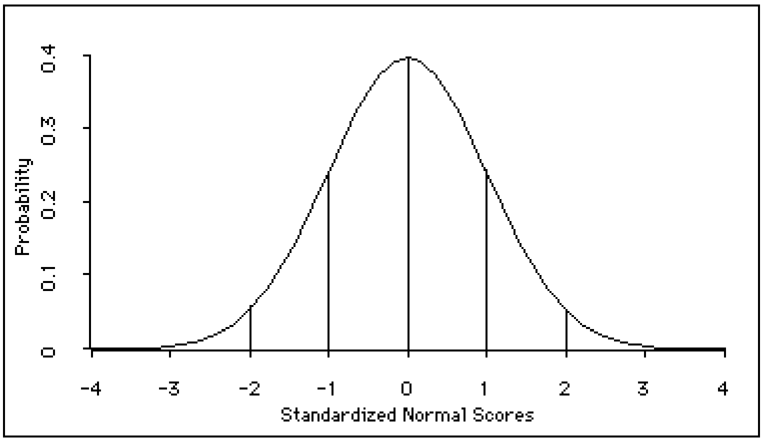
方法 1: 可以畫 histogram 或是 stem-leaf plot

– 若是呈現 bell-shaped, 則常態分配假設合理

方法 2: 可以畫 normal probability plot → 特殊的 Quantile-Quantile plot

– 若是呈現直線, 則常態分配假設合理

標準常態分配圖

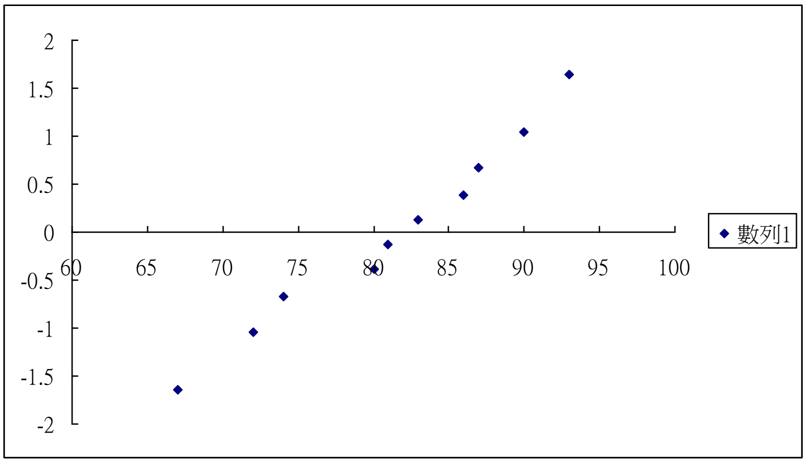


Calculations to obtain the normal plot → first step: order the data

i	$x_{(i)}$ 排序	$\frac{i}{n}$ (data PR)	$\frac{i-0.5}{n}$	$z_{(i)}$ (normal PR) (由查表得)
1	67	0.1	0.05	-1.645
2	72	0.2	0.15	-1.04
3	74	0.3	0.25	-0.675
4	80	0.4	0.35	-0.385
5	81	0.5	0.45	-0.125
6	83	0.6	0.55	0.125
7	86	0.7	0.65	0.385
8	87	0.8	0.75	0.675
9	90	0.9	0.85	1.04
10	93	1	0.95	1.645

Remarks:

- 如果 data PR 與 (standard) normal PR 兩者呈一直線,則 normal 假設合理



- 注意上圖橫軸與縱軸的尺度不一樣,使得角度看起來是 45% (其實不是).

註:

- QQ plot 的意義: 45 度線
  - 縱軸 U = $N(\mu,\sigma^2)$ PR
  - 橫軸 V = data PR
- Normal plot: 直線, 但不一定 45 度
  - 縱軸 Z = $N(0,1)$ PR
  - 橫軸 V = data PR

兩者關係:

如果 $U=V$ (原45度線) $\underset{轉換後}{\rightarrow} Z=\frac{U-\mu}{\sigma}=\frac{-\mu}{\sigma}$(截距)$+\frac{1}{\sigma}V$(非45度斜率)

補充一: Multiple comparison → 重要, but 不考

問題: 如果要 test $\mu_1 = \mu_2 = ... = \mu_k$，可不可以做 $\binom{k}{2}$ 組的 “兩兩比較”?

問題: 如果要 test $\mu_1 = \mu_2 \& \theta_1 = \theta_2 \& \tau_1 = \tau_2$ (數個參數一起比), 可不可以做 $\binom{3}{2}$ 組的 “兩兩比較”?

第一類問題: One-way factorial design: 單因子實驗與分析

概述: 探討 “溫度” (factor) 對 “細菌數目” (response)

實驗設計: 設定 K 個溫度的 “level” 進行實驗, 將 n 個培養皿 隨機分配給 K 組. 如果每組的個數相等, 則每個 treatment group 有 n/K 個人.

(註: 每組個數相同稱為 balanced design, 不同稱為 unbalanced design).

想要 test $H_0 : \mu_1 = \mu_2 = ... = \mu_K$

→ 若 K 組<u>變異數相同</u>, 用 ANOVA (Analysis of Variance)

統計量的型式: 和“組間差異” 與 “組內差異” 的比值有關.

第二類或是一般問題: 欲合併兩兩比較結果, 需要調整 type-I error

Suppose we want to test several hypotheses at a time.

Type-one error 的計算

a. A single test

α = Pr(reject $H_0 | H_0$ is correct)

= 1 - Pr(accept $H_0 | H_0$ is correct) = 0.05

b. Two tests: $H_0^{(1)}$ and $H_0^{(2)}$ 一起檢測

example: $H_0^{(1)} : \mu_1 = \mu_{10}$ and $H_0^{(2)} : \mu_2 = \mu_{20}$

（DNA 上同時看兩個位置是否符合假說）

換言之, $H_0 = H_0^{(1)} \cap H_0^{(2)}$（兩件事都要對, 才能接受 H_0）

令 α = Pr(reject $H_0^{(1)} | H_0^{(1)}$ is correct)

= Pr(reject $H_0^{(2)} | H_0^{(2)}$ is correct) = 0.05.

（個別的檢定都以 0.05 做為顯著水準）

For testing the hypothesis $H_0 = H_0^{(1)} \cap H_0^{(2)}$, the type one error is

α^* = 1 - Pr(accept $H_0 | H_0$ is correct)

= 1 - Pr(accept $H_0 = H_0^{(1)} \cap H_0^{(2)} | H_0^{(1)}, H_0^{(2)}$ are both correct)

≈ 1 - Pr(accept $H_0^{(1)} | H_0^{(1)}$ is correct)* Pr(accept $H_0^{(2)} | H_0^{(2)}$ is correct)

= 1 – 0.95*0.95 = 0.0975 > 0.05

c. M tests: $H_0^{(1)}, H_0^{(2)}, \ldots H_0^{(M)}$ 一起做.

（例如 DNA 上看 M 個位置是否都符合假說）

If the type one error rate for a single test is α, the type one error for the composite test approximately equals $1-(1-\alpha)^M >> \alpha$.

調整的方法 - Bonferroni adjustment:

- Set α^* = Pr(reject any of $H_0^{(1)},\ldots,H_0^{(M)} | H_0^{(1)},\ldots,H_0^{(M)}$ are all true)

  通常比 0.05 大

- Set α = Pr(accept $H_0^{(j)} | H_0^{(j)}$ is correct) = α^* / M

  → 單一檢定的 type I error

- Let $\alpha^* = 0.1, M = 10, \alpha = 0.1/10 = 0.01$

  可得 $1-(1-0.01)^{10} \approx 0.9$

註：還存在其它調整方法，都是出於不能直接合併數個 two-sample tests，所提出的修正

補充: Inference of proportion - one sample →不考

$$X = \sum_{i=1}^{n} Y_i \sim Binomial(n, p), \text{ where } Y_i \sim^{iid} Bernoulli(p)$$

$$\Pr(Y_i = y) = p^y (1-p)^{1-y} \quad \text{for} \quad y = 0,1, \quad E(Y_i) = p; \quad Var(Y_i) = p(1-p) = pq$$

Inference for p

Data: (X, n) Estimate of p: $\hat{p} = X / n$

Large-sample distribution theory: (中央極限定理)

$$\hat{p} = X / n = \sum_{i=1}^{n} Y_i / n \sim_{n\to\infty} N(p, Var(\hat{p}) = \frac{p(1-p)}{n})$$

$$\frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} \sim_{n\to\infty} N(0,1) \quad \textbf{or} \quad \frac{X - np}{\sqrt{np(1-p)}} \sim_{n\to\infty} N(0,1)$$

Proportion 推論的挑戰:

- $Var(\hat{p}) = p(1-p)$ 代表“標準化”過程牽涉到“欲推論的參數”(p)
- 過去參數為 μ，標準化除以 $\sqrt{Var(\bar{X})} = \sigma / \sqrt{n}$ (另一個參數), 問題單純

信賴區間: $(1-\alpha)100\%$ **Confidence interval for** p

$$\hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \quad \text{(標準差用估計的)}$$

Exercise: 民調最小樣本數的計算: 欲達到 $\pm 3\%$ 的誤差範圍, n 至少要多大?

$$1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \le 0.03 \;\rightarrow\; 1.96^2 \frac{10000}{9} \hat{p}\,(1-\hat{p}) \le n$$ (p 要選多少, 才能求樣本數)

民調大多用 p=0.5 (最保守的算法, 樣本數要一千多人)

p	1-p	n
0.5	0.5	1067.111
0.1	0.9	384.16
0.01	0.99	42.2576

這次調查於四月廿六日至五月十四日晚間進行，共計成功訪問了一萬七千二百三十二位成年人，另三千三百人拒訪；除連江縣完訪四百六十六人、金門縣六百人、澎湖縣六百零三人外，其他縣市有效樣本在七百至七百四十九人之間；在百分之九十五的信心水準下，各縣市抽樣誤差在正負三點六至四點六個百分點以內。

Hypothesis testing for p **(以下公式的標準化帶入估計的亦可)**

$$z_{obs} = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

Case 1: $H_0 : p = p_0$ versus $H_a : p > p_0$

Method 1: Reject H_0 if $z_{obs} > z_\alpha$

Method 2: p-value = $\Pr(Z > z_{obs})$

Case 2: $H_0 : p = p_0$ versus $H_a : p < p_0$

Method 1: Reject H_0 if $z_{obs} < -z_\alpha$

Method 2: p-value = $\Pr(Z < z_{obs})$

Case 3: $H_0 : p = p_0$ versus $H_a : p \neq p_0$

Method 1: Reject H_0 if $|z_{obs}| > z_{\alpha/2}$

Method 2: p-value = $2\Pr(Z > |z_{obs}|)$

Method 3: check whether $p_0 \in \hat{p} \pm z_{\alpha/2}\sqrt{\frac{p_0(1-p_0)}{n}}$

Topic: Two sample study – comparing two proportions → 不考

Data: (X_1, n_1) and (X_2, n_2)

Sample 1: $X_1 = \sum_{i=1}^{n_1} Y_{1i} \sim Binomial(n_2, p_1)$, where $Y_{1i} \sim^{iid} Bernoulli(p_1)$

Sample 2: $X_2 = \sum_{i=1}^{n_2} Y_{2i} \sim Binomial(n_2, p_2)$, where $Y_{2i} \sim^{iid} Bernoulli(p_2)$

Estimate of p_j**:** $\hat{p}_j = X_j / n_j \quad j = 1,2$

Large-sample distribution theory: (中央極限定理)

$$\frac{\hat{p}_1 - \hat{p}_2 - (p_1 - p_2)}{\sqrt{Var(\hat{p}_1) + Var(\hat{p}_2)}} = \frac{\hat{p}_1 - \hat{p}_2 - (p_1 - p_2)}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}} \sim_{n_j \to \infty} N(0,1)$$

$(1-\alpha)100\%$ **Confidence interval for** $p_1 - p_2$

$$(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} \quad \text{(標準差用估計的)}$$

Hypothesis testing for $p_1 - p_2$

Under $H_0: p_1 = p_2 = p$,

$$\frac{\hat{p}_1 - \hat{p}_2 - 0}{\sqrt{\frac{p(1-p)}{n_1} + \frac{p(1-p)}{n_2}}} \sim_{n_j \to \infty} N(0,1)$$

One can estimate p by $\hat{p} = \dfrac{X_1 + X_2}{n_1 + n_2}$ (用於標準化後的分母)

$n_1 + n_2 =$ 合併後的總樣本數; $X_1 + X_2$ = 合併後的總成功數

$$\hat{V}ar\left(\hat{p}\right) = \frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2} \underset{H_0}{=} \frac{\hat{p}(1-\hat{p})}{n_1} + \frac{\hat{p}(1-\hat{p})}{n_2} = \hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)$$

$$z_{obs} = \frac{\left(\hat{p}_1 - \hat{p}_2\right) - 0}{\sqrt{\hat{p}(1-\hat{p})(\frac{1}{n_1} + \frac{1}{n_2})}}$$

Case 1: $H_0: p_1 - p_2 = 0$ versus $H_a: p_1 - p_2 > 0$

Method 1: Reject H_0 if $z_{obs} > z_\alpha$

Method 2: p-value = $\Pr(Z > z_{obs})$

Case 2: $H_0: p_1 - p_2 = 0$ versus $H_a: p_1 - p_2 < 0$

Method 1: Reject H_0 if $z_{obs} < -z_\alpha$

Method 2: p-value = $\Pr(Z < z_{obs})$

Case 3: $H_0: p_1 - p_2 = 0$ versus $H_a: p_1 - p_2 \neq 0$

Method 1: Reject H_0 if $|z_{obs}| > z_{\alpha/2}$

Method 2: p-value = $2\Pr(Z > |z_{obs}|)$

Method 3: check whether

$$0 \in (\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})(\frac{1}{n_1} + \frac{1}{n_2})}$$

Remarks: 應用只要記得（1）大樣本才可用;（2）標準差估計要特別注意即可

補充：統計量的分配（以下不在考試範圍）

背景：

- 上學期我們學過隨機變數，與重要的分配（上學期 Lectures 4-7）

  $X \sim$ 分配，$\Pr(X = x)$（或是 $f(x)$），計算 $E(X), Var(X), \ldots$

- 隨機樣本 $X_i \overset{iid}{\sim}$ 某分配 for $i = 1, \ldots, n$

- <u>Statistics</u>（統計量）: functions of the sample (functions of the data)

  用來估計就叫做估計量：$\bar{X} = \sum_{i=1}^{n} X_i / n$，$\hat{\beta} = r \dfrac{S_Y}{S_X}$ …

  用來檢定就稱檢定統計量：$T = \dfrac{\bar{X} - \mu_0}{S_X / \sqrt{n}}$

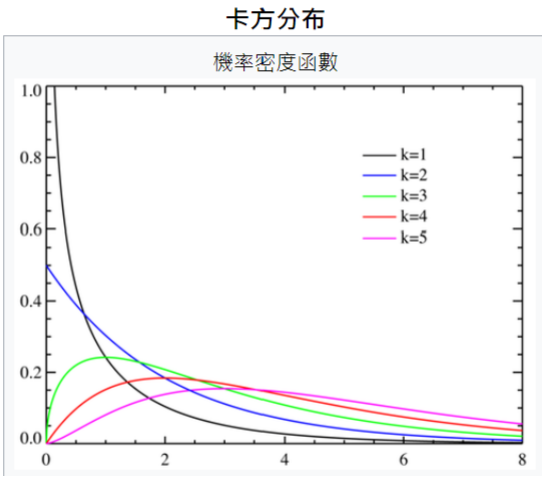
- 只要是隨機變數（不論是自然觀測的隨機變數，或是把一些隨機變數做加減乘除 … 的統計輛），都有其分配，後者叫做抽樣分配 (sampling distribution)

常見的抽樣分配

(I). 卡方分配 (Chi-squared distribution): 和“常態分配的平方” 有關

其定義：$\mathrm{X}_K^2 = (Z_1)^2 + \ldots + (Z_K)^2$ - Chi-squared distribution with K degree of freedom, where $Z_i \overset{iid}{\sim} N(0,1)(i = 1, \ldots, K)$

變異數估計量通常是”平方和”的型式，故往往與卡方分配有關



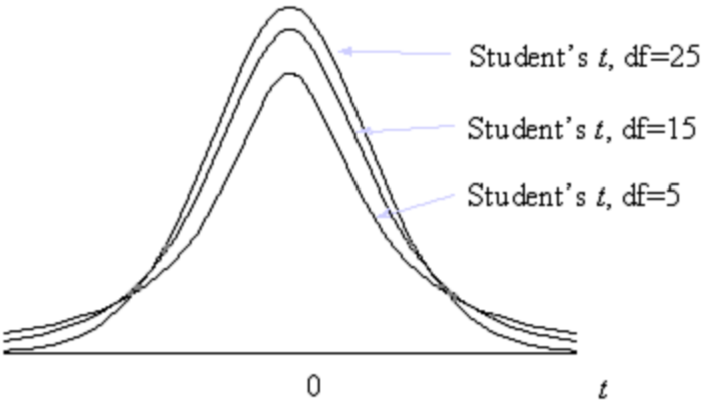
(II). T 分配 (t distribution): “常態除以開根號的卡方” → t 分配

numerator: normal, denominator: Chi-squared

Student's t is the distribution with n degrees of freedom of :

$$\frac{z}{\sqrt{\chi^2/n}}$$

- where z is the standard normal variable and χ^2 is a chi-square random variable with n degrees of freedom.



(III) F 分配 → 兩個彼此獨立的卡方相除可得 F 分配

$$F = \frac{X_a^2/d_1}{X_b^2/d_2} \sim F_{d_1,d_2},$$

$\mathrm{X}_{d_1}^2$ ~ Chi-squared distribution with degree of freedom d_1

$\mathrm{X}_{d_2}^2$ ~ Chi-squared distribution with degree of freedom d_2

$\mathrm{X}_{d_1}^2$ and $\mathrm{X}_{d_2}^2$ are independent

The F-distribution with d_1 and d_2 degrees of freedom is the distribution of

$$X = \frac{S_1/d_1}{S_2/d_2}$$

where S_1 and S_2 are independent random variables with chi-square distributions with respective degrees of freedom d_1 and d_2.

補充： Inference about σ^2 - One sample

Population: $X_i \sim^{iid} N(\mu,\sigma^2)$ - (μ,σ^2) are both unknown

Interest: σ^2

Objectives:

- Confidence interval for σ^2
- Hypothesis testing for σ^2

Procedures of developing the inference formula

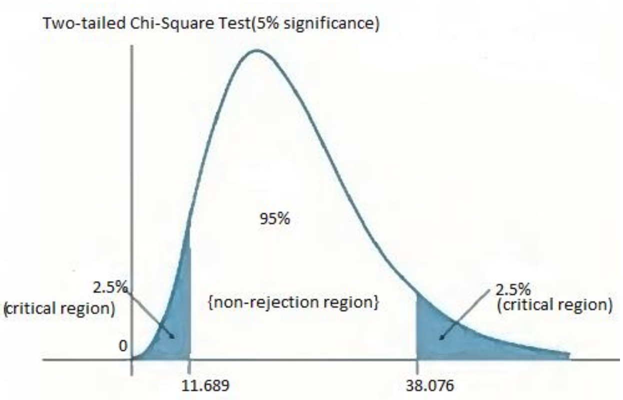
1. Point estimator of σ^2

Sample variance: $S^2 = \sum_{i=1}^{n}(X_i - \bar{X})^2/(n-1)$

2. Distribution theory of S^2

When $X_i \sim^{iid} N(\mu,\sigma^2)$,

$\frac{(n-1)S^2}{\sigma^2} \sim$ Chi-squared distribution with degree of freedom $n-1$



3. $(1-\alpha)100\%$ **confidence interval for σ^2**

$$\Pr(x^2_{1-\alpha/2,n-1} < \frac{(n-1)S^2}{\sigma^2} < x^2_{\alpha/2,n-1}) = 1-\alpha$$

$$\Pr(\frac{(n-1)S^2}{x^2_{\alpha/2,n-1}} < \sigma^2 < \frac{(n-1)S^2}{x^2_{1-\alpha/2,n-1}}) = 1-\alpha,$$

where $\Pr(X^2_{n-1} > x^2_{\alpha/2,n-1}) = \alpha/2$ and $\Pr(X^2_{n-1} > x^2_{1-\alpha/2,n-1}) = 1-\alpha/2$.

（注意卡方分配並不對稱，兩個切點的值並沒有明顯關係）
上次的卡方表資訊不夠（只有右尾的機率），附上另一個課本的卡方表

Formula of $(1-\alpha)100\%$ confidence interval: $(\frac{(n-1)S^2}{x^2_{\alpha/2,n-1}}, \frac{(n-1)S^2}{x^2_{1-\alpha/2,n-1}})$

補充： $(1-\alpha)100\%$ **upper confidence interval for σ^2**

$$\Pr(x^2_{1-\alpha,n-1} < \frac{(n-1)S^2}{\sigma^2}) = 1-\alpha \rightarrow \sigma^2 < \frac{(n-1)S^2}{x^2_{1-\alpha,n-1}}$$

Example: 4.63 $n = 51$， $s = 0.37$， $x^2_{0.025,50} =$ 71.42 $x^2_{0.975,50} =$ 32.36

$$\left(\frac{(n-1)S^2}{X^2_{\alpha/2,n-1}}, \frac{(n-1)S^2}{X^2_{1-\alpha/2,n-1}}\right) = \left(\frac{50\times(0.37^2)}{71.42}, \frac{50\times(0.37^2)}{32.36}\right)$$

4. Hypothesis testing for σ^2

$H_0 : \sigma^2 = \sigma_0^2$ vs. H_a (3 種情形如下)

Case 1: $H_a : \sigma^2 > \sigma_0^2$:

Method 1: Reject if $\dfrac{(n-1)S^2}{\sigma_0^2} > x^2_{\alpha,n-1}$

Method 2: p-value = $\Pr(X^2_{n-1} > \dfrac{(n-1)s^2}{\sigma_0^2})$

Example: $H_0 : \sigma = 0.35$ vs. $H_a : \sigma > 0.35$

Method 1: Reject if $\dfrac{50\times(0.37^2)}{(0.35)^2} > \; x^2_{0.05,50} = 67.50$

Method 2: p-value = $\Pr(X^2_{50} > \dfrac{50\times 0.37^2}{0.35^2}) = \Pr(X^2_{50} > 55.88) \quad \in (0.1, 0.5)$

Case 2: $H_a : \sigma^2 < \sigma_0^2$

Method 1: Reject if $\dfrac{(n-1)S^2}{\sigma_0^2} < x^2_{1-\alpha,n-1}$

Method 2: p-value = $\Pr(X^2_{n-1} < \dfrac{(n-1)s^2}{\sigma_0^2})$

Case 3: $H_a : \sigma^2 \neq \sigma_0^2$

Method 1: Reject if $\dfrac{(n-1)S^2}{\sigma_0^2} > x^2_{\alpha/2,n-1}$ or $\dfrac{(n-1)S^2}{\sigma_0^2} < x^2_{1-\alpha/2,n-1}$

Method 2: p-value: find the one-sided tail probability and then double it!!

Method 3: check whether $\sigma_0^2 \in \left(\dfrac{(n-1)S^2}{x^2_{\alpha/2,n-1}}, \dfrac{(n-1)S^2}{x^2_{1-\alpha/2,n-1}}\right)$

Example: $H_0 : \sigma = 0.35$ vs. $H_a : \sigma^2 \neq 0.35$

p-value = $2\Pr(X^2_{50} > \dfrac{50\times 0.37^2}{0.35^2}) = 2\Pr(X^2_{50} > 55.88) \quad \in (0.2, 1)$

Since 0.1225 falls in (0.096,0.212), accept $H_0: \sigma = 0.35$

補充: Comparison between σ_1^2 and σ_2^2 - two samples

Population 1: $X_{1i} \sim^{iid} N(\mu_1, \sigma_1^2)$ $(i = 1, ..., n_1)$

Population 2: $X_{2j} \sim^{iid} N(\mu_2, \sigma_2^2)$ $(j = 1, ..., n_2)$

Interest: $\dfrac{\sigma_1^2}{\sigma_2^2}$

Note:

- **for the mean:** $\mu_1 - \mu_2$
- 比較可以基於“差別”或是“比值”

Objectives:

- Confidence interval for $\dfrac{\sigma_1^2}{\sigma_2^2}$
- Hypothesis testing for $\dfrac{\sigma_1^2}{\sigma_2^2}$

Procedures of developing the inference formula

1. **Point estimator of** $\dfrac{\sigma_1^2}{\sigma_2^2}$

   Sample variance: $S_1^2 = \sum_{i=1}^{n_1} (X_{1i} - \bar{X}_1)^2 / (n_1 - 1)$; $S_2^2 = \sum_{j=1}^{n_1} (X_{2j} - \bar{X}_2)^2 / (n_2 - 1)$

2. **Distribution theory of** S^2

   When the two populations are both normally distributed,

$$F = \frac{S_1^2 / \sigma_1^2}{S_2^2 / \sigma_2^2} \sim F_{n_1 - 1, n_2 - 1}$$

(F distribution with degrees of freedom $(n_1 - 1, n_2 - 1)$)

Note: F 分配有兩個自由度