12/04統計學實習課

114-1統計學實習課

第三次小考檢討

114-1統計學實習課

A.Multiple Choice (60%)

1. We use CLT to estimate the parameters, means that the distribution of the _____ which has a sufficient large numbers of independent random variables, each with finite mean and variance, will be approximately normally distributed. a. mean of population b. population c. mean of sample d. sample

- **答案: c. mean of sample**
- **解析:** 中央極限定理 (CLT) 指出，當樣本數 (n) 足夠大時，樣本平均數 ($\bar{X}$, mean of sample) 的抽樣分佈會近似於常態分佈，無論母體分佈為何。

A.Multiple Choice (60%)

2. A simple random sample of 100 observations was taken from a large population. The sample mean and the standard deviation were determined to be 80 and 12 respectively. The standard error of the mean is a. 1.20 b. 0.12 c. 8.00 d. 0.80

- **答案: a. 1.20**
- **解析:** 標準誤 (Standard Error, SE) 的公式為 $s/\sqrt{n}$。 $SE = 12/\sqrt{100} = 12/10 = 1.20$。

A.Multiple Choice (60%)

3. Stratified random sampling is a method of selecting a sample in which a. the sample is first divided into strata, and then random samples are taken from each stratum b. various strata are selected from the sample c. the population is first divided into strata, and then random samples are drawn from each stratum d. None of these alternatives is correct.

- **答案: c. the population is first divided into strata, and then random samples are drawn from each stratum**
- **解析:** 分層隨機抽樣 (Stratified random sampling) 的定義是先將「母體 (population)」分成不同的層 (strata)，再從每一層中隨機抽取樣本。選項 a 說是將「樣本」分層，這是錯誤的。

A.Multiple Choice (60%)

4. Random samples of size 36 are taken from an infinite population whose mean and standard deviation are 20 and 15, respectively. The distribution of the population is unknown. The mean & the standard error of the sample mean are a. 36 and 15 b. 20 and 15 c. 20 and 0.417 d. 20 and 2.5

- **答案: d. 20 and 2.5**
- **解析:**
  - 樣本平均數的期望值 (Mean of sample mean) 等於母體平均數 $\mu = 20$。
  - 樣本平均數的標準誤 (Standard error) $\sigma_{\bar{x}} = \sigma/\sqrt{n} = 15/\sqrt{36} = 15/6 = 2.5$。

A.Multiple Choice (60%)

5. As the sample size increases, the variability among the sample means a. increases b. decreases c. remains the same d. depends upon the specific population being sampled

- **答案: b. decreases**
- **解析:** 樣本平均數的變異程度由標準誤 ($\sigma/\sqrt{n}$) 決定。當樣本數 n 變大，分母變大，整體標準誤變小，因此變異性降低。

A.Multiple Choice (60%)

6. From a population that is not normally distributed and whose standard deviation is not known, a sample of 6 items is selected to develop an interval estimate for the mean of the population (μ). a. The normal distribution can be used. b. The t distribution with 5 degrees of freedom must be used. c. The t distribution with 6 degrees of freedom must be used. d. The sample size must be increased.

- **答案: d. The sample size must be increased.**
- **解析:** 母體非常態且樣本數極小 ($n = 6$)。在這種情況下，既不能用 Z 分佈（因為非常態且 n 小），也不能用 t 分佈（t 分佈理論上假設母體為常態）。為了使用中央極限定理來進行估計，必須增加樣本數 ($n \geq 30$)。

A.Multiple Choice (60%)

7. A 95% confidence interval for a population mean is determined to be 100 to 120. If the confidence coefficient is reduced to 0.90, the interval for μ a. becomes narrower b. becomes wider c. does not change d. becomes 0.1

- **答案: a. becomes narrower**
- **解析:** 信賴水準 (Confidence Level) 降低（從 95% 降到 90%），對應的關鍵值 (Z 或 t score) 會變小，導致誤差範圍 (Margin of Error) 變小，因此區間寬度會變窄。

A.Multiple Choice (60%)

8. A machine that produces a major part for an airplane engine is monitored closely. In the past, 10% of the parts produced would be defective. With a .95 probability, the sample size that needs to be taken if the desired margin of error is .04 or less is a. 110 b. 111 c. 216 d. 217

- **答案: d. 217**
- **解析:** 樣本數公式 $n = p(1-p)(Z/E)^2$。

  $p = 0.10$

  $E = 0.04$

  95% 信賴水準對應 $Z \approx 1.96$

  $n = 0.1 \times 0.9 \times (1.96/0.04)^2$

  $n = 0.09 \times (49)^2 = 0.09 \times 2401 = 216.09$

  樣本數必須無條件進位，故為 217。

A.Multiple Choice (60%)

9. An interval estimate is a range of values used to estimate a. the shape of the population's distribution b. the sampling distribution c. a sample statistic d. a population parameter

- **答案: d. a population parameter**
- **解析:** 區間估計（如信賴區間）是用一個數值範圍來估計未知的「母體參數」（如母體平均數 μ 或比例 p）。

A.Multiple Choice (60%)

10. The population is normally distributed and whose mean and standard deviation is μ and σ respectively. (X_1, X_2, X_3) is a random sample from the population. There are four estimators: $T_1 = (3X_1 + 3X_2 + 4X_3)/10; T_2 = (X_1 + X_2 + X_3)/3; T_3 = (X_1 + 2X_2 + 3X_3)/6; T_4 = (2X_1 + 3X_2 + 4X_3)/9$. Which is an unbiased estimator and has the smallest variance? a. T_1 b. T_2 c. T_3 d. T_4

- **答案: b.** T_2
- **解析:**
  1. **不偏性檢查 (Unbiasedness):** 係數總和必須為 1。
     - $T_1 : (3 + 3 + 4)/10 = 1$ (不偏)
     - $T_2 : (1 + 1 + 1)/3 = 1$ (不偏)
     - $T_3 : (1 + 2 + 3)/6 = 1$ (不偏)
     - $T_4 : (2 + 3 + 4)/9 = 1$ (不偏)

A.Multiple Choice (60%)

2. **變異數比較 (Variance):** $Var(aX) = a^2 Var(X)$。我們比較係數平方和。
   - $Var(T_1) = [(3^2 + 3^2 + 4^2)/100]\sigma^2 = (34/100)\sigma^2 = 0.34\sigma^2$
   - $Var(T_2) = [(1^2 + 1^2 + 1^2)/9]\sigma^2 = (3/9)\sigma^2 = 0.333\sigma^2$ (最小)
   - $Var(T_3) = [(1^2 + 2^2 + 3^2)/36]\sigma^2 = (14/36)\sigma^2 \approx 0.389\sigma^2$
   - $Var(T_4) = [(2^2 + 3^2 + 4^2)/81]\sigma^2 = (29/81)\sigma^2 \approx 0.358\sigma^2$
- T_2 (即樣本平均數) 擁有最小變異數。

A.Multiple Choice (60%)

11. The power curve provides the probability of a. correctly accepting the null hypothesis b. incorrectly accepting the null hypothesis c. correctly rejecting the alternative hypothesis d. correctly rejecting the null hypothesis

- **答案: d. correctly rejecting the null hypothesis**
- **解析:** 檢定力 (Power, $1 - \beta$) 的定義是：當虛無假設 (H_0) 為假時，正確拒絕它的機率。

A.Multiple Choice (60%)

12. Which of the following does not need to be known in order to compute the p-value? a. knowledge of whether the test is one-tailed or two-tailed b. the value of the test statistic c. the level of significance d. None of these alternatives is correct.

- **答案: c. the level of significance**
- **解析:** P-value 是根據檢定統計量 (test statistic) 和檢定方向 (one/two-tailed) 算出來的機率值。它是一個客觀數值，不需要知道顯著水準 α (level of significance) 就能計算出來。(α 是用來與 P-value 比較以做結論的標準，但不是計算 P-value 的參數)。

A.Multiple Choice (60%)

13. A random sample of 100 people was taken. Eighty-five of the people in the sample favored Candidate A. We are interested in determining whether or not the proportion of the population in favor of Candidate A is significantly more than 80%. At 95% confidence, it can be concluded that the proportion of the population in favor of candidate A a. is significantly greater than 80% b. is not significantly greater than 80% c. is significantly greater than 85% d. is not significantly greater than 85%

- **答案: b. is not significantly greater than 80%**
- **解析:**

$$H_0 : p \leq 0.80, H_1 : p > 0.80$$

$$\hat{p} = 0.85, n = 100, p_0 = 0.80$$

$$Z = \frac{0.85 - 0.80}{\sqrt{\frac{0.80 \times 0.20}{100}}} = \frac{0.05}{0.04} = 1.25$$

95% 單尾檢定的臨界值 $Z_{0.05} = 1.645$。

因為 $1.25 < 1.645$，未落在拒絕域，故無法拒絕 H_0，結論是沒有顯著大於 80%。

A.Multiple Choice (60%)

14. In hypothesis testing, the tentative assumption about the population parameter is a. the alternative hypothesis. b. the null hypothesis. c. either the null or the alternative. d. neither the null nor the alternative.

- **答案: b. the null hypothesis.**
- **解析:** 在假設檢定中，我們暫時假設虛無假設 (H_0) 是真的，直到有足夠證據推翻它。

A.Multiple Choice (60%)

15. For a one-tailed test (lower tail), a sample size of 10 at 90% confidence, t = a. 1.383 b. 2.821 c. -1.383 d. -2.821

- **答案: c. -1.383**
- **解析:**

  $n = 10$, 自由度 $df = 9$。

  90% confidence (對於單尾檢定通常指 $\alpha = 0.10$)。

  查 t 表 $t_{0.10,9} = 1.383$。

  因為是左尾檢定 (lower tail)，臨界值為負數，故為 -1.383。

B. Subjective part (45%)

1. A university planner is interested in determining the percentage of spring semester students who will attend summer school. She takes a pilot sample of 160 spring semester students discovering that 56 will return to summer school.

(a) Construct a 95% confidence interval estimate for the percentage of spring semester students who will return to summer school.

- **計算:**

  樣本比例 $\hat{p} = \frac{56}{160} = 0.35$

  樣本數 $n = 160$

  95% 信賴水準對應 $Z = 1.96$

  標準誤 $SE = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{0.35 \times 0.65}{160}} = \sqrt{\frac{0.2275}{160}} \approx 0.0377$

  誤差範圍 (Margin of Error) $ME = 1.96 \times 0.0377 \approx 0.0739$

  信賴區間 = 0.35 ± 0.0739

- **答案: [0.2761, 0.4239]** (或 27.61% 到 42.39%)

B. Subjective part (45%)

(b) Using the results of the pilot study with a 0.95 probability, how large of a sample would have to be taken to provide a margin of error of 3% or less?

- **計算:**

  $$E = 0.03$$

  $$Z = 1.96$$

  使用 pilot study 的 $\hat{p} = 0.35$ 作為規劃值。

  $$n = \frac{(Z)^2 \hat{p}(1-\hat{p})}{E^2} = \frac{(1.96)^2 \times 0.35 \times 0.65}{(0.03)^2}$$

  $$n = \frac{3.8416 \times 0.2275}{0.0009} \approx \frac{0.873964}{0.0009} \approx 971.07$$

  樣本數需無條件進位。

- **答案: 972**

B. Subjective part (45%)

(c) Is the return percentage less than 0.3 significantly? ($\alpha = 0.05$, sample size is 160)

- **計算:**

  $H_0 : p \geq 0.30$

  $H_1 : p < 0.30$ (左尾檢定)

  $\hat{p} = 0.35$

  由於樣本比例 (0.35) 實際上大於假設比例 (0.30)，我們不可能得到顯著「小於」的結論。

  驗證計算: $Z = \frac{0.35 - 0.30}{\sqrt{\frac{0.3 \times 0.7}{160}}} = \frac{0.05}{0.0362} \approx 1.38$

  左尾檢定 $\alpha = 0.05$ 的拒絕域是 $Z < -1.645$。計算出的 $Z = 1.38$ 不在拒絕域內。

- **答案: No, it is not significantly less than 0.3.** (因為樣本比例反而比 0.3 高)

B. Subjective part (45%)

2. An experimental diet to induce weight loss was followed for one week by a randomly selected group of 12 students with the following results.

Student	1	2	3	4	5	6	7	8	9	10	11	12
Loss in Pounds	2.2	2.6	0.4	2.0	0.0	1.8	5.2	3.8	4.2	3.8	1.4	2.6

(a) Find a point estimate for the average amount lost after one week on this diet. Is this an unbiased estimate of the population mean? Explain.

- **計算:**

  總和 $\sum x = 2.2 + 2.6 + 0.4 + 2.0 + 0.0 + 1.8 + 5.2 + 3.8 + 4.2 + 3.8 + 1.4 + 2.6 = 30$

  樣本平均 $\bar{x} = 30/12 = 2.5$

- **答案:** Point estimate is **2.5**.
- **解釋:** Yes, it is unbiased. The expected value of the sample mean is equal to the population mean ($E[\bar{x}] = \mu$). (樣本平均數的期望值等於母體平均數)。

B. Subjective part (45%)

(b) Find a point estimate for the variance of the amount lost on this diet. Is this an unbiased estimate of the population variance? Explain.

- **計算:**

  樣本變異數 $s^2 = \frac{\sum(x_i - \bar{x})^2}{n-1}$

  計算各數值平方和 $\sum x^2 = 2.2^2 + \ldots + 2.6^2 = 101.28$

  $SS = \sum x^2 - \frac{(\sum x)^2}{n} = 101.28 - \frac{30^2}{12} = 101.28 - 75 = 26.28$

  $s^2 = \frac{26.28}{11} \approx 2.389$

- **答案:** Point estimate is **2.389**.
- **解釋:** Yes, it is unbiased. The sample variance formula uses $n - 1$ in the denominator specifically to make it an unbiased estimator of σ^2 ($E[s^2] = \sigma^2$).

B. Subjective part (45%)

(c) Find a point estimate for the standard deviation of the amount lost on this diet.

- **計算:**
  - $s = \sqrt{s^2} = \sqrt{2.3891}$
- **答案: 1.546** (約 1.55)
- *註：雖然 s^2 是不偏估計量，但 s 其實是 σ 的有偏估計量，但通常作為點估計值使用。*

B. Subjective part (45%)

3. A local university administers a comprehensive examination to the candidates for B.S. degrees in Business Administration. Five examinations are selected at random and scored. The scores are shown below. Scores: 80, 90, 91, 62, 77

B. Subjective part (45%)

(a) Compute the mean and the standard deviation of the sample.

- **計算:**

  Mean $\bar{x} = \frac{80+90+91+62+77}{5} = \frac{400}{5} = 80$

  Variance s^2:

  $$(80-80)^2 = 0$$

  $$(90-80)^2 = 100$$

  $$(91-80)^2 = 121$$

  $$(62-80)^2 = (-18)^2 = 324$$

  $$(77-80)^2 = (-3)^2 = 9$$

  Sum of squares = $0 + 100 + 121 + 324 + 9 = 554$

  $$s^2 = 554/(5-1) = 554/4 = 138.5$$

  Std Dev $s = \sqrt{138.5} \approx 11.7686$

- **答案: Mean = 80, Standard Deviation = 11.77**

B. Subjective part (45%)

(b) Compute the margin of error at 95% confidence. Assume the population is normally distributed.

- **計算:**

$$n = 5, df = 4$$

查 t 表 $t_{0.025,4} = 2.776$

$$ME = t \times \frac{s}{\sqrt{n}} = 2.776 \times \frac{11.7686}{\sqrt{5}} = 2.776 \times 5.263 \approx 14.61$$

- **答案: 14.61**

B. Subjective part (45%)

(c) Develop a 95% confidence interval estimate for the mean of the population.

- **計算:**

  CI = $\bar{x} \pm ME$

  80 ± 14.61

  Lower limit: $80 - 14.61 = 65.39$

  Upper limit: $80 + 14.61 = 94.61$

- **答案: [65.39, 94.61]**