8. 基本抽樣分佈

Fundamental Sampling Distributions

- 平均值的抽樣分佈與中央極限定理
- S^2的抽樣分佈
- t-分佈
- F-分佈

8.1 平均值的抽樣分佈與中央極限定理

定義 8.1

統計值(量)的機率分佈稱為抽樣分佈(sampling distribution)。

定理 8.1 中央極限定理 (Central Limit Theorem)

假設隨機變數 $\bar{X}$ 代表一從平均值為 μ，變異數為 σ^2 的母體中所選取之大小為 n 的隨機樣本的平均值。

如果 $n \to \infty$，則隨機變數 $Z = \dfrac{\bar{X} - \mu}{\sigma / \sqrt{n}}$

n >= 30

n < 30

為一標準常態分佈 $n(z\,;0\,,1)$。

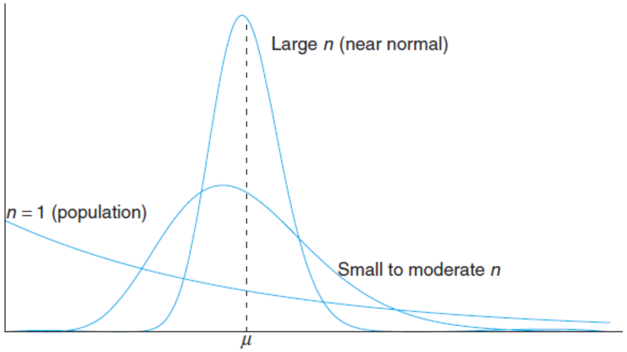


Illustration of the Central Limit Theorem
(distribution of $\bar{X}$ for $n = 1$, moderate n, and large n).

範例8.1

- An electrical firm manufactures light bulbs that have a length of life that is approximately normally distributed, with mean equal to 800 hours and a standard deviation of 40 hours. Find the probability that a random sample of 16 bulbs will have an average life of less than 775 hours.

Case Study 8.1

- An important manufacturing process produces cylindrical component parts for the automotive industry. It is important that the process produce parts having a mean diameter of 5.0 millimeters. The engineer involved conjectures that the population mean is 5.0 millimeters. An experiment is conducted in which 100 parts produced by the process are selected randomly and the diameter measured on each. It is known that the population standard deviation is $\sigma = 0.1$ millimeter. The experiment indicates a sample average diameter of $\bar{x} = 5.027$ millimeters. Does this sample information appear to support or refute the engineer's conjecture?

範例8.2

- Traveling between two campuses of a university in a city via shuttle bus takes, on average, 28 minutes with a standard deviation of 5 minutes. In a given week, a bus transported passengers 40 times. What is the probability that the average transport time was more than 30 minutes? Assume the mean time is measured **to the nearest minute**.

8.1 平均值的抽樣分佈與中央極限定理(contd.)

定理 8.2

已知有 2 個母體，其平均值分別為 μ_1 與 μ_2，變異數分別為 σ_1^2 與 σ_2^2。如果大小分別為 n_1 個樣本與 n_2 個樣本是從 2 個母體隨機且獨立挑選出來的，若變數 $\bar{X}_1 - \bar{X}_2$ 代表其平均值的差，則 $\bar{X}_1 - \bar{X}_2$ 的抽樣分佈近似於常態分佈，

且其平均值 $\mu_{\bar{X}_1 - \bar{X}_2} = \mu_1 - \mu_2$，變異數 $\sigma^2_{\bar{X}_1 - \bar{X}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$，因此 $Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$ 為近似於標準常態分佈的隨機變數。

Case Study 8.2

- Two independent experiments are run in which two different types of paint are compared. Eighteen specimens are painted using type A, and the drying time, in hours, is recorded for each. The same is done with type B. The population standard deviations are both known to be 1.0. Assuming that the mean drying time is equal for the two types of paint, find $P(\bar{X}_A - \bar{X}_B > 1.0)$, where $\bar{X}_A$ and $\bar{X}_B$ are average drying times for samples of size $n_A = n_B = 18$.

Case Study 8.2 (contd.)

- Suppose that the experiment is conducted for the purpose of drawing an inference regarding the equality of the two population mean drying times, μ_A and μ_B. What inference will you make?

範例8.3

- The television picture tubes of manufacturer A have a mean lifetime of 6.5 years and a standard deviation of 0.9 year, while those of manufacturer B have a mean lifetime of 6.0 years and a standard deviation of 0.8 year. What is the probability that a random sample of 36 tubes from manufacturer A will have a mean lifetime that is at least 1 year more than the mean lifetime of a sample of 49 tubes from manufacturer B?

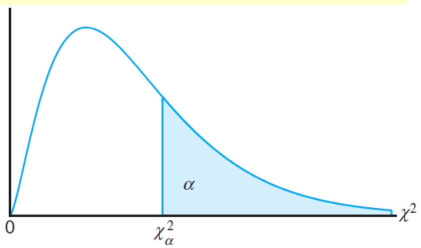
8.2 S^2的抽樣分佈

定理 8.3

如果 S^2 是從一平均值為 μ，變異數為 σ^2 的常態分佈中隨機選取之大小為 n 的隨機樣本的變異數，則統計值

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2} = \sum_{i=1}^{n} \frac{(X_i - \bar{X})^2}{\sigma^2}$$

是一開方分佈且其自由度為 $\upsilon = n-1$



定理 8.4

$$S^2 = \frac{1}{n(n-1)}\left[n\sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i\right)^2\right]$$

How to prove it?

範例8.4

- A manufacturer of car batteries guarantees that the batteries will last, on average, 3 years with a standard deviation of 1 year. If five of these batteries have lifetimes of 1.9, 2.4, 3.0, 3.5, and 4.2 years, should the manufacturer still be convinced that the batteries have a standard deviation of 1 year? Assume that the battery lifetime follows a normal distribution.

$\sigma^2 = 1$

8.3 t－分佈

定理 8.5

令 Z 為標準常態隨機變數且 V 為自由度為 υ 之開方隨機變數。如果 Z 以及 V 是獨立的，則隨機變數 $T=\dfrac{Z}{\sqrt{V/\upsilon}}$ 的密度函數為

$$h(t)=\frac{\Gamma[(\upsilon+1)/2]}{\Gamma(\upsilon/2)\sqrt{\pi\upsilon}}\left(1+\frac{t^2}{\upsilon}\right)^{-(\upsilon+1)/2},\quad -\infty<t<\infty$$

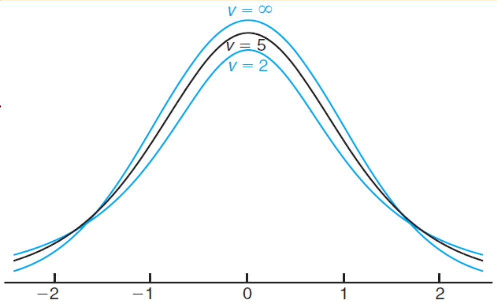
此密度函數稱為自由度為 υ 的 t-分佈。

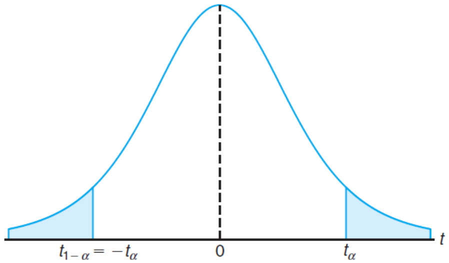
推論7.1

令 $X_1, X_2, \cdots, X_n$ 是獨立的隨機變數，且其機率分佈都是平均值為 μ，標準差為 σ 之常態分佈。令 $\bar{X}=\dfrac{1}{n}\sum_{i=1}^{n}X_i$ $S^2=\dfrac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$，則隨機變數 $T=\dfrac{\bar{X}-\mu}{S/\sqrt{n}}$ 是自由度為 $\upsilon=n-1$ 的 t-分佈。

範例8.5

- With $\upsilon=14$ (degrees of freedom), find the t-value that leaves (1) an area of 0.025 to the left, and (2) an area of 0.975 to the right.
- Find $P(-t_{0.025}<T<t_{0.05})$





範例8.6

- Find k such that P(k < T < −1.761) = 0.045 for a random sample of size 15 selected from a normal distribution.

範例8.7

- A chemical engineer claims that the population mean yield of a certain batch process is 500 grams per milliliter of raw material. To check this claim he samples 25 batches each month. If the computed t-value falls between $-t_{0.05}$ and $t_{0.05}$, he is satisfied with this claim. What conclusion should he draw from a sample that has a mean $\bar{x} = 518$ grams per milliliter and a sample standard deviation s = 40 grams? Assume the distribution of yields to be approximately normal.

8.4 F-分佈

定理 8.6

假設兩獨立隨機變數 U 和 V，已知 U, V 是開方分佈，其自由度分別為 υ_1 以及 υ_2。則隨機變數 $F=\dfrac{U/\upsilon_1}{V/\upsilon_2}$ 的密度函數為

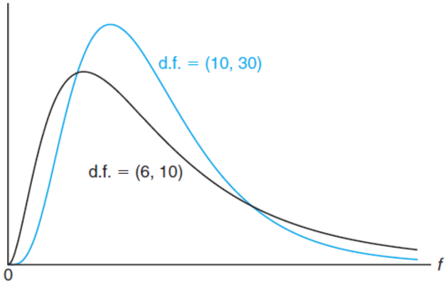
$$h(f)=\begin{cases}\dfrac{\Gamma[(\upsilon_1+\upsilon_2)/2](\upsilon_1/\upsilon_2)^{\upsilon_1/2}}{\Gamma(\upsilon_1/2)\Gamma(\upsilon_2/2)}\dfrac{f^{(\upsilon_1/2)-1}}{(1+\upsilon_1 f/\upsilon_2)^{(\upsilon_1+\upsilon_2)/2}}, & f>0\\ 0, & f\le 0\end{cases}$$

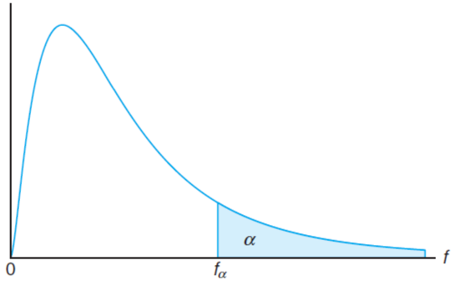
此函數稱為自由度為 υ_1 以及 υ_2 之F-分佈 (F-distribution)。

定理 8.7

若將自由度為 υ_1 以及 υ_2 之 f_α 值記為 $f_\alpha(\upsilon_1,\upsilon_2)$，則

$$f_{1-\alpha}(\upsilon_1,\upsilon_2)=\frac{1}{f_\alpha(\upsilon_2,\upsilon_1)}$$





$f_{0.95}(6,10)=?$

8.4 F-分佈

定理 8.8

S_1^2 以及 S_2^2 是樣本數為 n_1 以及 n_2 的 2 個獨立隨機樣本的變異數，此 2 個隨機樣本分別挑選自變異數為 σ_1^2 以及 σ_2^2 的常態母體，則

$$F = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} = \frac{\sigma_2^2 S_1^2}{\sigma_1^2 S_2^2}$$

是自由度為 $\upsilon_1 = n_1 - 1$ 以及 $\upsilon_2 = n_2 - 1$ 的 F-分佈。

範例8.8

- If S_1^2 and S_2^2 represent the variances of independent random samples of size n1 = 8 and n2 = 12, taken from normal populations with equal variances, find P(S_1^2 / S_2^2 < 4.89).