第一章　敘述統計學（考古題）

2006 年 10 月 18 日 最後修改

1.1（95-淡江-產經）

1. Answering the questions below for the following data set:
   2.0, 1.1, -1.5, 0.3, 1.9, -2.3, -1.2, 0.7, -0.5, 3.1, -1.5
   -0.2, 1.0, -0.6, 1.3, -0.9, -0.1, 0.8, -1.1, 2.4, 0.9, -0.1,
   a) Do the data appear to be bell-shaped? (6 分)
   b) Calculate the coefficient of skewness. (6 分)
   c) Using the empirical rule, estimate the range of values within which about 68% of the data value are expected to lie. (6 分)

【解】

x	x^2	x^3	x^4
2.0	4.00	8.000	16.0000
1.1	1.21	1.331	1.4641
-1.5	2.25	-3.375	5.0625
0.3	0.09	0.027	0.0081
1.9	3.61	6.859	13.0321
-2.3	5.29	-12.167	27.9841
-1.2	1.44	-1.728	2.0736
0.7	0.49	0.343	0.2401
-0.5	0.25	-0.125	0.0625
3.1	9.61	29.791	92.3521
-1.5	2.25	-3.375	5.0625
-0.2	0.04	-0.008	0.0016
1.0	1.00	1.000	1.0000
-0.6	0.36	-0.216	0.1296
1.3	1.69	2.197	2.8561
-0.9	0.81	-0.729	0.6561
-0.1	0.01	-0.001	0.0001
0.8	0.64	0.512	0.4096
-1.1	1.21	-1.331	1.4641
2.4	5.76	13.824	33.1776
0.9	0.81	0.729	0.6561
-0.1	0.01	-0.001	0.0001
5.5	42.83	41.557	203.6927

$$\mu = \frac{\Sigma x}{n} = \frac{5.5}{22} = 0.25, \qquad M_e = \frac{x_{11} + x_{12}}{2} = \frac{-0.1 + 0.3}{2} = 0.1$$

$$M_2 = \frac{\Sigma x^2 - n \times \mu^2}{n} = \frac{42.83 - 22 \times 0.25^2}{22} = 1.8843, \qquad \sigma = \sqrt{M_2} = 1.3727$$

$$M_3 = \frac{\Sigma x^3 - 3\mu\Sigma x^2 + 2n\mu^3}{n} = 0.4601, \qquad \alpha_3 = \frac{M_3}{\sigma^3} = 0.1779$$

$$M_4 = \frac{\Sigma x^4 - 4\mu\Sigma x^3 + 6\mu^2\Sigma x^2 - 3n\mu^4}{n} = 8.0881, \qquad \alpha_4 = \frac{M_4}{\sigma^4} = 2.2779$$

(a) $\alpha_3 = 0.1779 > 0$ ，$\alpha_4 = 2.2779 < 3$ ，因此這是一個右偏的低闊峰，不是常態峰。

(b) 偏態係數 $S_k=\dfrac{3\left(\mu-M_e\right)}{\sigma}=\dfrac{3\left(0.25-0.1\right)}{1.3727}=0.3278$

(c)

$$P\left(\mu-\sigma\le x\le\mu+\sigma\right)\approx 68\%$$
$$\Rightarrow\quad 0.25-1.3727\le x\le 0.25+1.3727$$
$$\Rightarrow\quad -1.1227\le x\le 1.6227$$

1.2（95-淡江-國貿）

> 1) Please explain the following terms:
> a) The Central Limit Theorem (中央極限定理). (10%)
> b) Chebyshev's Inequality (謝比雪夫不等式). (5%)
> c) The empirical rule (經驗法則). (5%)

【解】

(a)

$\bar{x}=\dfrac{x_1+x_2+\cdots+x_n}{n}$，其期望值、標準差分別爲 $\mu_{\bar{x}}$ 、 $s_{\bar{x}}$ ，則當 $n\to\infty$ 時，

$\bar{x}$ 成爲期望值 $\mu_{\bar{x}}$ 、標準差 $s_{\bar{x}}$ 的常態分配。

（關鍵字眼爲： $\bar{x}$ 、 $n\to\infty$ 、常態分配。）

(b)

對任意分配之 x，$P\left(\left|x-\mu\right|\le k\sigma\right)\ge 1-\dfrac{1}{k^2}$

(c)

鐘形分配之 x，$P\left(\left|x-\mu\right|\le\sigma\right)\doteq 68\%$ 、 $P\left(\left|x-\mu\right|\le 2\sigma\right)\doteq 95\%$ 、 $P\left(\left|x-\mu\right|\le 3\sigma\right)\doteq 99.7\%$

1.3（95-淡江-財金）

> 2. (20%)某班有 60 人，期中考平均成績(μ) 爲 75 分，標準差(σ) 爲 5 分，考慮下列二種情形：
> (1)若資料分配近似對稱分配，問全班考 60 分~90 分有幾人；不及格有幾人。
> (2)若資料分配不爲對稱分配，則全班考 60 分~90 分有幾人；不及格有幾人。

【解】

(1)用經驗法則

$z=\dfrac{\left|60-75\right|}{5}=\dfrac{\left|90-75\right|}{5}=3$ ， $P\left(60\le x\le 90\right)=99.7\%$ ， $n_{60\sim 90}=60\times 99.7\%\doteq 60$

約 60 人分數介於 60~90，沒有人不及格。

(2)用柴比雪夫定理

$$k=\frac{|60-75|}{5}=\frac{|90-75|}{5}=3\text{ ， }P(60\le x\le 90)=1-\tfrac{1}{3^2}=\tfrac{8}{9}\text{ ， }n_{60\sim 90}=60\times\tfrac{8}{9}\doteq 53.3$$

$$P(x<60)=\tfrac{1}{2}\times\tfrac{1}{3^2}=\tfrac{1}{18}\text{ ， }n_{x<60}=60\times\tfrac{1}{18}\doteq 3.3$$

約 53 人分數介於 60~90，約 3 人不及格。

1.4（95-雲科大-企管）

三、 某大都市內蔬果平均零售價格為萵苣每單位 71 元，SD= 5，蕃茄每單位 40 元，SD=3，菜瓜每單位 19 元，SD=2，假設此市內某間連鎖超市所販賣的價格為萵苣每單位 78 元，蕃茄 45 元，菜瓜 21 元，請問此三者的售價相對於市內平均價格而言高低順序為何？ 10%

(1) 萵苣＞蕃茄＞菜瓜(2) 蕃茄＞萵苣＞菜瓜(3) 萵苣＞菜瓜＞蕃茄
(4) 蕃茄＞菜瓜＞萵苣(5) 菜瓜＞萵苣＞蕃茄(6) 菜瓜＞蕃茄＞萵苣
(7) 萵苣= 蕃茄 =菜瓜

【解】（10/18/2006 修正）

計算三者的 z 值，結果爲

$$z_{\text{萵苣}}=\frac{78-71}{5}=1.40 \qquad z_{\text{蕃茄}}=\frac{45-40}{3}=1.67 \qquad z_{\text{菜瓜}}=\frac{21-19}{2}=1.00$$

因 $z_{\text{蕃茄}}>z_{\text{萵苣}}>z_{\text{菜瓜}}$ ，故選(2)。

1.5（94-政大-財管）

1. (20 points) True or False (Give a brief explanation.)

(A). Mean is always greater than median.

(B). For a linear program to have an optimal feasible solution, the number of constraints always has to be greater than the number of decision variables.

(C). A statistic $T(X)$ is a sufficient statistic for μ if the conditional distribution of the sample X given the value of $T(X)$ does not depend on μ.

(D). If Y1 and Y2 are disjoint events, then they are independent.

(E). $f(x)=\begin{cases}0 & x<0\\ x/2 & 0\le x<2\\ 0 & x\ge 2\end{cases}$ The mean of x is 8/3.

【解】

(a) False，右偏時才是。

(b) False，作業研究的內容。

(c) True，第五章估計的內容。

(d) False，

disjoint $\Rightarrow$ $Y_1 \cap Y_2 = \varnothing$ $\Rightarrow$ 與獨立與否無關

(e) False，第三章的內容。

$$E(x) = \int xf(x)dx = \int_0^2 x\frac{x}{2}dx = \frac{1}{2}\frac{x^3}{3}\Big|_0^2 = \frac{4}{3}$$

1.6（94-政大-財管）

5. (10 points) Suppose there are 10 people earning 1,000 respectively, and there is 1 person earning 10,000 by himself. The total income earned by these 11 people is 20,000. The mean income is 1,818.182. The first 10 people earn less than the mean income and consider themselves low-income class.

A. Why does this situation happen?

B. Is "mean" a good statistics at this case? If not, what else statistics can we use?

【解】

(A) 因爲有一個收入佔總收入一半的『離群值』。

(B) 有離群值時平均數不是一個恰當的集中趨勢衡量指標，這時可以考慮用中位數。

1.7（94-政大-風管）

9. The actual proportion of families in a certain city who own, rather than rent, their home is 0.70. If 84 families in this city are interviewed at random and their responses to the question whether or not they own their home are looked upon as values of independent random variables having identical Bernoulli distributions with parameter $p = 0.70$, with what probability can we assert that the value we obtain for the sample proportion $\hat{p}$ will fall between 0.64 and 0.76, using the result of Chebyshev's theorem and the central limit theorem? **(10 points)**

【解】

$$\sigma = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.7 \times 0.3}{84}} = 0.05$$

(a)柴比雪夫定理

$$k=\frac{|0.76-0.7|}{0.05}=1.2 \quad \Rightarrow \quad P(0.64\le \hat{p}\le 0.76)\ge 1-\frac{1}{1.2^2}=\frac{11}{36}=0.3056$$

(b)常態分配

$$z=\frac{|0.76-0.7|}{0.05}=1.2 \quad \Rightarrow \quad P(0.64\le \hat{p}\le 0.76)=P(-1.2\le z\le 1.2)=0.7699$$

1.8（93-雲科大-企管）

1. 某校研究所入學考試統計學共 340 人應試，該科考生平均成績為 72，標準差為 8。若這次成績分配近似常態，請估計：
   (1) 在 64 分與 80 分間約有幾位考生？__________(5%)
   (2) 在 56 分與 88 分間約有幾位考生？__________(5%)

【解】

(1) 1 個標準差之內，機率為 68%， $n=340\times 68\%=231.2\doteq 231$

(2) 2 個標準差之內，機率為 95%， $n=340\times 95\%=323$

1.9（93-政大-財管）

1. (10%) According to Chebyshev's theorem, at least what proportion of the data will be within $\mu \pm k\sigma$ for each value of k?
   a. k=2
   b. k=3.2

【解】

(a) $1-1/2^2=\frac{3}{4}$

(b) $1-\frac{1}{3.2^2}=0.9023$

1.10（93-政大-財管）

2. (10%) Financial analysts like to use the standard deviation as a measure of risk for a stock. The greater the deviation in a stock price over time, the more risky it is to invest in the stock. However, the average prices of some stocks are considerably higher than the average price of others, allowing for the potential of a greater standard deviation of price. For example, a standard deviation of $5.00 on a $10.00 stock is considerably different from a $5.00 standard deviation on a $40.00 stock. In this situation, a coefficient of variation might provide insight into risk. Suppose stock X costs an average of $32.00 per share and has had a standard deviation of $3.45 for the past 60 days. Suppose stock Y costs an average of $84.00 per share and has had a standard deviation of $5.40 for the past 60 days. Use the coefficient of variation to determine the variability for each stock.

【解】

$$CV_x = \frac{\sigma_x}{\mu_x} \times 100\% = \frac{3.45}{32} \times 100\% = 10.78\%$$

$$CV_y = \frac{\sigma_y}{\mu_y} \times 100\% = \frac{5.40}{84} \times 100\% = 6.43\%$$

1.11（93-台大-商學）

2. Each person who applies for an assembly job at North Carolina Furniture is given a mechanical aptitude test. One part of the test involves assembling a dresser based on numbered instructions. A sample of the lengths of time it took 42 persons to assemble the dresser was organized into the following frequency distribution.

Length of time (minutes)	Number
2 up to 4	4
4 up to 6	8
6 up to 8	14
8 up to 10	9
10 up to 12	5
12 up to 14	2

What is the variance?

(A) 4.1402 (B) 4.8579 (C) 5.5376 (D) 6.7387 (E) None of the above.

【解】

時間	次數	代表值	x*f	x²*f
2 up to 4	4	3	12	36
4 up to 6	8	5	40	200
6 up to 8	14	7	98	686
8 up to 10	9	9	81	729
10 up to 12	5	11	55	605
12 up to 14	2	13	26	338
	42		312	2594

$$s^2 = \frac{\Sigma x^2 - (\Sigma x)^2 / n}{n-1} = \frac{2594 - 312^2/42}{42-1} = 6.7387$$