

考試科目	統計學	所別	統計學系	考試時間	月 星期	日 上午	第 節
------	-----	----	------	------	---------	---------	--------

1. A Statistics exam is given to the students at three departments in a school. The scores on the exam at each department are regarded as having normal distributions with unknown means μ_1 , μ_2 , and μ_3 , respectively, and unknown common variance σ^2 ($\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma^2$). Using the data in the accompanying table on independent random samples from each department, test to see if evidence exists of a difference between μ_1 and μ_2 . Use $\alpha = 0.05$. (10 pts)

Department I	Department II	Department III
$n_1 = 10$	$n_2 = 10$	$n_3 = 10$
$\sum x_i^2 = 36,950$	$\sum y_i^2 = 25,850$	$\sum w_i^2 = 49,900$
$\bar{x} = 60$	$\bar{y} = 50$	$\bar{w} = 70$

2. A random sample of 5,726 telephone numbers taken in March 2000 yielded 1,105 which were unlisted, and one year later a sample of 5,384 yielded 980 unlisted numbers.
- Test at level 0.1 to see if there is a difference in true proportions of unlisted numbers between the two years. (6 pts)
 - Construct a 90% confidence interval for the difference in true proportions of unlisted numbers between the two years. (4 pts)
3. Let Y_1, \dots, Y_{100} be independent random variables, each having a normal distribution with mean 10 and standard deviation 5.
- Since the random variables are identically distributed,

$$P(8 < Y_1 < 15) = P(8 < Y_2 < 15) = \dots = P(8 < Y_{100} < 15).$$
 Find the probability. (5%)
 - Let X count the number of the Y_i 's between 8 and 15. What is the distribution of X? Why? (6%)
 - Find $P(10 \leq X \leq 59)$, where X is as in (b). You may want to use an approximation. (5%)
 - Find the probability that all 100 random variables (Y_1, \dots, Y_{100}) take values between 8 and 15. (4%)

考試科目	統計學	所別	統計學系	考試時間	12月2日 13:30-15:10
------	-----	----	------	------	-------------------

4. 探索性資料分析(Exploratory Data Analysis, 簡稱 EDA)是資料分析的第一步,也是非常重要的第一步,可以從 EDA 發掘重要訊息。以下是民國 94 年大學指定考試各科成績的基本統計量,包括中位數、平均數、標準差、四分位數等敘述性統計量:

	國文	英文	數學甲	數學乙	化學	物理	生物	歷史	地理
Min.	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.00
12%	27.00	8.00	11.00	4.00	8.00	6.00	22.00	13.0	18.00
1st Qu.	34.00	16.00	22.00	12.00	15.00	12.00	32.00	28.0	30.00
Median	44.00	34.00	34.00	29.00	34.00	23.00	45.00	39.0	39.00
Mean	43.56	36.68	36.36	34.36	38.88	28.75	46.16	38.7	39.51
3rd Qu.	53.00	56.00	49.00	56.00	60.00	41.00	60.00	50.0	49.00
88%	60.00	69.00	59.00	61.00	76.00	57.00	71.00	56.0	55.00
Max.	93.00	98.00	100.00	100.00	100.00	100.00	99.00	89.0	90.00
st.d.	13.88	23.88	18.72	25.97	27.00	21.50	19.39	16.20	14.46

- (a) 試以這些統計量判斷哪些科目的成績分佈較接近常態分配,哪些科目較像是右偏(Right skewed)分配,請分別舉出三個像是常態分配、三個像是右偏分配的科目。本題請詳述你/妳如何判斷常態分配、右偏分配,並說明這些判斷方法的依據。(10分)
- (b) 離群值(Outlier)一般的定義為與眾不同的觀察值,請以上述的基本統計量判斷哪些科目有離群值,請說明你/妳的判斷方法,並舉出三個存在離群值的科目。(6分)
- (c) 基本統計量有時也可用於判斷平均數的大小,請舉出可分辨出平均數不同的兩個科目,並說明你的判斷方法。(4分)

考試科目	統計學	所別	統計學系	考試時間	12月2日 13:30-15:10
------	-----	----	------	------	-------------------

5. A supervising inspector of incoming quality only wants to know whether there is any difference among the average lifetime of battery under the three choice temperatures, but it seems that different plate materials may also influence the variability of lifetime. The experiment is designed by a statistician and the observations are as follows.

Temp. (i)	plate material (j)				Total
	1	2	3	4	
100	21	28	25	32	106
125	17	26	21	23	87
150	10	4	6	8	28
	Total				221

Note: total sum of square $(SST = \sum_i \sum_j (y_{ij} - \bar{y}_{..})^2 = 954.92)$

- 5-1 What is the name for the designed experiment? How to perform the all experiments? (2%+3%)
- 5-2 What are the effect model of ANOVA and assumptions? (4%)
- 5-3 Derive the distributions of $y_{ij} - \bar{y}_{.j}$, $i=1,2,3, j=1,2,3,4$. (4%)
- 5-4 Is temp. an important factor for the lifetime? $\alpha = 0.05$ (5%)
- 5-5 What is the appropriate design of the experiment for the same problem in the future (2%) (Hint: the answer can be obtained from the result in 5-4)

6. (20 points) Suppose X_1, \dots, X_n is a sample from a $N(\mu, \sigma^2)$ population and consider the sample standard deviation $\hat{\sigma}$ and the sample-mean deviation $\bar{\sigma}$ defined by

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

and

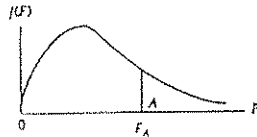
$$\bar{\sigma} = \frac{\sum_{i=1}^n |X_i - \bar{X}|}{n}$$

- (a) What is the distribution of $\sum_{i=1}^n (X_i - \bar{X})^2$? (5 points)
- (b) Find $E(\hat{\sigma}^2 - \sigma^2)^2$. (5 points)
- (c) What is the distribution of $X_i - \bar{X}$? (5 points)
- (d) Find $E(\bar{\sigma})$. (5 points)

考試科目	統計學	所別	統計學系	考試時間	12月2日 13:30-15:10
------	-----	----	------	------	-------------------

TABLE V THE F DISTRIBUTION TABLE

Percentage Points of the F Distribution, $\alpha = .05$



$\nu_2 \backslash \nu_1$	NUMERATOR DEGREES OF FREEDOM								
	1	2	3	4	5	6	7	8	9
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88