科目名稱	統計學	類組代碼	B11
		科目碼	B1192

※本項考試依簡章規定所有考科均「不可」使用計算機。　　本科試題共計 **7** 頁

考試題型及配分：選擇題（單選題）100%

1. The scores of a sample of 100 students are given below.
   mean = 60, range = 30, mode = 70, variance = 36, median = 65.
   What is the coefficient of variation?　(3%)
   (A) 1.0%　　　(B) 8.33%　　　(C) 10.0%　　　(D) 20.0%　　　(E) 60.0%

2. Suppose we samples 10 students and recorded their scores on a quiz: five scored 15, two scored 20, and three scored 25. We then discover that an error was made and that a score recorded as 15 was actually 20. Which of the following sample numerical measures will change when we do our computations with the correction?　(3%)
   (A) Median and mode.　　(B) Mean and mode.　　(C) Mean and range.
   (D) Variance and range.　　(E) Median and variance.

3. Which of the following rule or graph does little help to roughly determine a range of 90% probability for a random sample with large sample size?　(3%)
   (A) Empirical rule.　　　(B) Chebyshev's rule.　　　(C) Histogram.
   (D) Box plot.　　　　(E) Scatter plot.

4. Let the possible values of a random variable X be 0, 1, 2, and 3, and let
   $P\{X=1\}= 3/8$, $P\{X=0\}=P\{X=3\} = 1/8$. The variance of $Y = 2X+1$ is　(4%)
   (A) 1.5　　　(B) 3　　　(C) 4　　　(D) 5　　　(E) 16

5. Suppose an inspector randomly selects two of the 10 spare parts in a box for testing. If the box contains exactly four defective parts. If the inspector selects without replacement, what is the probability that the inspector will find at least one defective part?　(4%)
   (A) 0.16　　(B) 0.25　　(C) 0.33　　(D) 0.36　　(E) 0.67

6. In a factory, production lines I, II, and III account for 20, 30, and 50% of the total output. If 2% of the products from production line I are defective, and the corresponding percentages for production line II and III are 2% and 1%, what is the probability that a defective product will come from production line II ?　(4%)
   (A) 0.15　　(B) 0.27　　(C) 0.33　　(D) 0.4　　(E) None of the above

7. When determining the sample size for estimating a proportion p under a given confidence level and sampling error, the closer to 0.5 p is the _____ the sample size required. (3%)
   (A) larger　　　　(B) smaller　　　　(C) Sample size is not affected.
   (D) The sample size cannot be determined.　　　(E) None of the above.

8. On the average, 1.6 customers per minute arrive at any one of the checkout counters of Sunny market. What type of probability distribution can be used to find out the probability that there will be no customers arriving at a checkout counter in 10 minutes?? (3%)
   (A) Poisson distribution.　　　　(B) hypergeometric distribution.
   (C) binomial distribution.　　　　(D) normal distribution.
   (E) exponential distribution.

9. Suppose X is a continuous random variable and density equal to
   $$f(x) = \begin{cases} ce^{-5x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$ Find the value of constant c. (3%)
   (A) -0.2　　　(B) 0.2　　　(C) 1　　　(D) -5　　　(E) 5

10. Consider a normal random variable with $\mu = 1$ and standard deviation $\sigma = 1$. Which of the following is true? (3%)
    (A) $P(X > 0.645) = 0.1$　　　　(B) $P(X < -1.96) = 0.05$
    (C) $P(X < 3) < 1 - P(X > -3)$　　　(D) $P(X < 1.5) = P(X > 0.5)$
    (E) $P(X = 0) \neq P(X = 1)$

11. Which of the following statement is correct about a 90% confidence interval of μ? (3%)
    (A) We are 90% confident that our sample mean equals the population mean μ.
    (B) If we repeatedly draw samples of the same size from the same population, 90% of the resulting confidence intervals will include μ.
    (C) There is a 90% probability that the population mean μ will lie between the lower confidence limit and the upper confidence limit.
    (D) 90% of the population values will lie within the confidence interval.
    (E) None of the above.

12. From a population with a variance of 100, a sample of 25 items is selected. At 95% confidence, the margin of error for estimating the population mean is? (3%)
    (A) 2　　　(B) 3.29　　　(C) 3.92　　　(D) 4　　　(E) 5

13. Which of the following does not need to be known in computing the p-value of testing the population mean? (3%)
    (A) Knowledge of whether the test is one-tailed or two-tailed.
    (B) The value of the test statistic.　　(C) Whether the population variance is known.

(D) The level of significance.　　　(E) None of the above.

14. If we want to reduce the probability of a Type II error to 0.1 in testing the hypothesis $H_0 : \mu \leq 12$ versus $H_a : \mu > 12$ when the population mean is 13 and the population variance is 16. What sample size is recommended? (α=0.05) (3%)
(A) 274　　　(B) 35　　　(C) 78　　　(D) 1068　　　(E) 137

15. A confidence interval for $(\mu_1 - \mu_2)$ is (3, 6). Which of the following inferences is true? (3%)
(A) $\mu_1 = \mu_2$　　　(B) $\mu_1 > \mu_2$　　　(C) No significant difference between means.
(D) $\mu_1 < \mu_2$　　　(E) We cannot decide it.

16. If $X_i, i = 1, ..., n$, from a population follows a normal distribution with unknown mean μ and variance $\sigma^2 = 25$. To construct the 95% confidence interval for population mean and maintain a margin of error within 0.5, what is the smallest required sample size? (3%)
(A) 168　　　(B) 196　　　(C) 283　　　(D) 385　　　(E) 400

17. The grades of a sample of 5 students, selected from a large population, are given below.
Grade: 65, 75, 55, 85, 70. At 5% significance level to test if the variance of the population is larger than 50.　Which of the following answer is correct? (4%)
(A) $H_0: \sigma^2 \geq 50$; $H_a: \sigma^2 < 50$, reject H_0.
(B) $H_0: \sigma^2 \leq 50$; $H_a: \sigma^2 > 50$, reject H_0.
(C) $H_0: \sigma^2 = 50$; $H_a: \sigma^2 \neq 50$, reject H_0.
(D) $H_0: \sigma^2 \geq 50$; $H_a: \sigma^2 < 50$, do not reject H_0.
(E) $H_0: \sigma^2 \leq 50$; $H_a: \sigma^2 > 50$, do not reject H_0.

18. An interval estimate is a range of values used to estimate? (3%)
(A) the shape of the population's distribution
(B) the sampling distribution　　　(C) a sample statistic
(D) the variation of the population　　　(E) a population parameter

19. Previous records show that the average time spent on the questionnaire was 6 minutes for participants of a study. A sample of 81 participants was randomly chosen, and their completion times recorded. The mean completion time was 5.5 minutes with a standard deviation of 2.7 minutes. In order to test whether the time to fill the questionnaire has been reduced, which of the following answer is correct? (α=0.05)　　　(4%)

(A) $H_0: \mu = 6; H_a: \mu \neq 6$, reject H_0. 　　(B) $H_0: \mu \leq 6; H_a: \mu > 6$, reject H_0.

(C) $H_0: \mu \geq 6; H_a: \mu < 6$, reject H_0. 　　(D) $H_0: \mu \leq 6; H_a: \mu > 6$, do not reject H_0.

(E) $H_0: \mu \geq 6; H_a: \mu < 6$, do not reject H_0.

20. A marketing study was conducted to compare the variation in the age of male and female purchasers of a certain product. Random and independent samples were selected for both male and female purchasers of the product. The sample data is shown here:

Female: n = 31, sample mean = 50.3, sample standard deviation = 13.3

Male: n = 21, sample mean = 39.8, sample standard deviation = 10.0

At 5% significance level to test if the variation σ_1^2 in the female ages exceeds the variation σ_2^2 in the male ages. Which of the following answer is correct? (4%)

(A) $H_0: \sigma_1^2 \geq \sigma_2^2; H_a: \sigma_1^2 < \sigma_2^2$, reject H_0. 　　(B) $H_0: \sigma_1^2 \leq \sigma_2^2; H_a: \sigma_1^2 > \sigma_2^2$, reject H_0.

(C) $H_0: \sigma_1^2 = \sigma_2^2; H_a: \sigma_1^2 \neq \sigma_2^2$, reject H_0.

(D) $H_0: \sigma_1^2 \geq \sigma_2^2; H_a: \sigma_1^2 < \sigma_2^2$, do not reject H_0.

(E) $H_0: \sigma_1^2 \leq \sigma_2^2; H_a: \sigma_1^2 > \sigma_2^2$, do not reject H_0.

21. In constructing a confidence interval estimate for the difference between two population means, we: (3%)

(A) pool the population variances when they are equal.

(B) pool the population variances when the population means are equal.

(C) pool the population variances when the populations are normally distributed.

(D) never pool the population variances. 　　(E) None of the above.

22. Which of the following is not a required assumption for the analysis of variance? (3%)

(A) The response variable for each population is normally distributed.

(B) The mean of the response variable is the same for each population.

(C) The observations must be independent.

(D) At least 2 populations are under consideration.

(E) Populations have equal variances.

23. If the experimental units are heterogeneous, an experimental design can be used to form homogeneous groups is (3%)

(A) Randomized design. 　　(B) Completely randomized design.

(C) Factorial design. 　　(D) Randomized block design. 　　(E) None of the above.

24. Consider the following table, the null hypothesis for this ANOVA problem is? (3%)

Source of Variation	Sum of Squares	DF	Mean Square	F
Between Treatments		3		3
Within Treatments			10	
Total		9		

(A) $\mu_1 = \mu_2 = \mu_3$　　　(B) $\mu_1 = \mu_2 = \mu_3 = \mu_4$　　　(C) $\mu_1 = \mu_2 = ... = \mu_6$

(D) $\sigma_1^2 = \sigma_2^2 = ... = \sigma_4^2$　　　(E) None of the above.

25. Refer to above problem, the total sum of squares (SST) is　(4%)

(A) 30　　(B) 90　　(C) 150　　(D) 240　　(E) None of the above.

26. The following data were obtained for a <u>randomized block design</u> involving 4 treatments and 3 blocks: SST = 440, SSBL = 80, SSE = 60. Use $\alpha = 0.05$ to test whether the treatment effects are zero or not. The test statistic is (4%)

(A) 4.0　　(B) 7.5　　(C) 10.0　　(D) 13.3　　(E) 15.5

27. The calculations for a factorial experiment involving 3 levels of factor A, 2 levels of factor B, and 3 replications resulted in the following data: SST=258, SSA=60, SSB=48, SSE=120. Use α=0.05 to test the interaction effect of factor A and factor B. The test statistic is ? (4%)

(A) 1.5　　(B) 3.0　　(C) 4.8　　(D) 10.0　　(E) 12.0

28. Of the following, the one that is minimized when the method of least squares is applied in a simple linear regression analysis is　(3%)

(A) $\sum(Y_i - \widehat{Y}_i)^2$　　　(B) $\sum(X_i - \widehat{X}_i)^2$　　　(C) $\sum(Y_i - \overline{Y})^2$　　　(D) $\sum(\widehat{Y}_i - \overline{Y})^2$

(E) $\sum(\widehat{X}_i - \overline{X})^2$

29. In a multiple linear regression, if all the points of a scatter diagram lie on the curvilinear $y = x^2 + 1$, the coefficient of determination (R^2) is?　(3%)

(A) $R^2 = 0$　　(B) $R^2 = 1$　　(C) $R^2 > 1$　　(D) $0 < R^2 < 1$

(E) None of above alternatives is correct.

30. A sample of 26 provided a simple linear regression $\widehat{y} = 1.2 - 0.8x$, SSR (sum of squares due to regression) = 24, and SSE (sum of squares due to error）= 144. What is the t statistic for testing the significance of the slope? (4%)

(A) 4　　(B) 2　　(C) -2　　(D) -4

(E) None of above alternatives is correct.

科目名稱	統計學	類組代碼	**B11**
		科目碼	**B1192**

※本項考試依簡章規定所有考科均「不可」使用計算機。　　本科試題共計　**7**　頁

※ χ^2 distribution, $P(\chi_r^2 > \chi_r^2(\alpha)) = \alpha$

$\chi_2^2(0.05) = 5.99$; $\chi_3^2(0.05) = 7.82$; $\chi_4^2(0.05) = 9.49$; $\chi_5^2(0.05) = 10.07$.

※ F distribution, $P(F_{n,m} > F_{n,m}(\alpha)) = \alpha$

$F_{30,20}(0.01) = 2.78$;　$F_{30,20}(0.025) = 2.35$;　$F_{30,20}(0.05) = 2.04$;　$F_{30,20}(0.1) = 1.74$

$F_{20,30}(0.01) = 2.55$;　$F_{20,30}(0.025) = 2.20$;　$F_{20,30}(0.05) = 1.93$;　$F_{20,30}(0.1) = 1.67$

Table S1　Cumulative Probabilities for the standard Normal Distribution

Number in the table represents $P(Z \le z)$
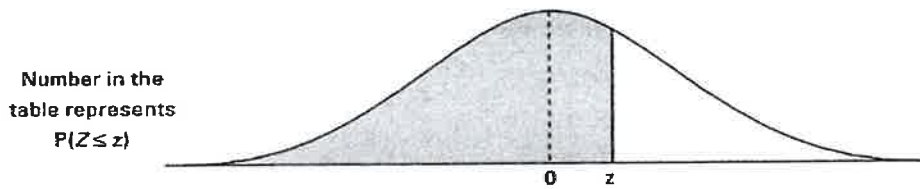
z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998
3.5	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998
3.6	.9998	.9998	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999

Table S2　　t Distribution

Numbers in each row of the table are values on a t-distribution with
(df) degrees of freedom for selected right-tail (greater-than) probabilities (p).
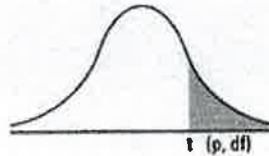
df/p	0.40	0.25	0.10	0.05	0.025	0.01	0.005	0.0005
1	0.324920	1.000000	3.077684	6.313752	12.70620	31.82052	63.65674	636.6192
2	0.288675	0.816497	1.885618	2.919986	4.30265	6.96456	9.92484	31.5991
3	0.276671	0.764892	1.637744	2.353363	3.18245	4.54070	5.84091	12.9240
4	0.270722	0.740697	1.533206	2.131847	2.77645	3.74695	4.60409	8.6103
5	0.267181	0.726687	1.475884	2.015048	2.57058	3.36493	4.03214	6.8688
6	0.264835	0.717558	1.439756	1.943180	2.44691	3.14267	3.70743	5.9588
7	0.263167	0.711142	1.414924	1.894579	2.36462	2.99795	3.49948	5.4079
8	0.261921	0.706387	1.396815	1.859548	2.30600	2.89646	3.35539	5.0413
9	0.260955	0.702722	1.383029	1.833113	2.26216	2.82144	3.24984	4.7809
10	0.260185	0.699812	1.372184	1.812461	2.22814	2.76377	3.16927	4.5869
11	0.259556	0.697445	1.363430	1.795885	2.20099	2.71808	3.10581	4.4370
12	0.259033	0.695483	1.356217	1.782288	2.17881	2.68100	3.05454	43178
13	0.258591	0.693829	1.350171	1.770933	2.16037	2.65031	3.01228	4.2208
14	0.258213	0.692417	1.345030	1.761310	2.14479	2.62449	2.97684	4.1405
15	0.257885	0.691197	1.340606	1.753050	2.13145	2.60248	2.94671	4.0728
16	0.257599	0.690132	1.336757	1.745884	2.11991	2.58349	2.92078	4.0150
17	0.257347	0.689195	1.333379	1.739607	2.10982	2.56693	2.89823	3.9651
18	0.257123	0.688364	1.330391	1.734064	2.10092	2.55238	2.87844	3.9216
19	0.256923	0.687621	1.327728	1.729133	2.09302	2.53948	2.86093	3.8834
20	0.256743	0.686954	1.325341	1.724718	2.08596	2.52798	2.84534	3.8495
21	0.256580	0.686352	1.323188	1.720743	2.07961	2.51765	2.83136	3.8193
22	0.256432	0.685805	1.321237	1.717144	2.07387	2.50832	2.81876	3.7921
23	0.256297	0.685306	1.319460	1.713872	2.06866	2.49987	2.80734	3.7676
24	0.256173	0.684850	1.317836	1.710882	2.06390	2.49216	2.79694	3.7454
25	0.256060	0.684430	1.316345	1.708141	2.05954	2.48511	2.78744	3.7251
26	0.255955	0.684043	1.314972	1.705618	2.05553	2.47863	2.77871	3.7066
27	0.255858	0.683685	1.313703	1.703288	2.05183	2.47266	2.77068	3.6896
28	0.255768	0.683353	1.312527	1.701131	2.04841	2.46714	2.76326	3.6739
29	0.255684	0.683044	1.311434	1.699127	2.04523	2.46202	2.75639	3.6594
30	0.255605	0.682756	1.310415	1.697261	2.04227	2.45726	2.75000	3.6460
z	0.253347	0.674490	1.281552	1.644854	1.95996	2.32635	2.57583	3.2905
CI	———	———	80%	90%	95%	98%	99%	99.9%