國立成功大學
115學年度碩士班招生考試試題

編　號：211

系　所：經濟學系

科　目：統計學

日　期：0203

節　次：第 1 節

注　意：1.不可使用計算機
　　　　2.請於答案卷(卡)作答，於試題上作答，不予計分。

Part I. Multiple Choice Questions (60 points)

Instructions: This section consists of 20 questions, each worth 3 points. Each question has **one and only one** correct answer. Choose the best option for each question. Your answers must be written in the **answer sheet** provided following the illustrated table format below. Mark your choice clearly and **do not** write your answers on this question paper. The standard normal distribution table is provided at the end of the exam.

Answer Sheet Illustration (Q1–Q20): For Reference Only. DO NOT Write Your Answers on This Question Paper!!

Q1		Q2		Q3		Q4		Q5	
Q6		Q7		Q8		Q9		Q10	
Q11		Q12		Q13		Q14		Q15	
Q16		Q17		Q18		Q19		Q20	

---

1. Let X and Z be random variables with finite and strictly positive variances, and suppose that $\mathrm{Cov}(X, Z) \neq 0$. Define a new random variable

$$Y = 3X - 5Z.$$

Which of the following statements is correct?

(A) The correlation between Y and X depends only on $\mathrm{Var}(X)$ and is independent of $\mathrm{Cov}(X, Z)$.

(B) The regression coefficient of Y on X equals 3 because Y is defined as a linear function of X.

(C) The correlation between Y and X must be equal to 1 since Y depends linearly on X.

(D) The regression coefficient of Y on X is

$$\beta_{Y,X} = 3 - 5\frac{\mathrm{Cov}(X,Z)}{\mathrm{Var}(X)}.$$

2. Suppose the joint density of (X, Y) is

$$f(x,y) = 6x^2y, \quad 0 < x < 1, \ 0 < y < 1.$$

Which statement is correct?

(A) X and Y are not independent and $E(X \mid Y)$ depends on Y

(B) X and Y are independent and $E(X \mid Y) = E(X)$

(C) X and Y are independent but $E(X \mid Y) \neq E(X)$

(D) X and Y are not independent but $E(X \mid Y) = E(X)$

3. Suppose annual earnings in an economy are normally distributed with mean \$50,000 and a positive standard deviation. A government defines the poverty line as the income level below which 5% of the population falls.

Which statement is correct?

(A) The poverty line is determined by overall income dispersion

(B) The poverty line is an inequality measure summarizing income differences

(C) The poverty line is independent of the earnings distribution

(D) The poverty line is defined as a cutoff based on the lower tail of the earnings distribution

4. An applied econometrician studies an outcome variable Y and two information sets $\mathcal{I}_1 \subset \mathcal{I}_2$, where $\mathcal{I}_1$ represents baseline information and $\mathcal{I}_2$ includes additional covariates. Assume all expectations below are well-defined.

Which statement must always hold?

(A) $\mathbb{E}[\mathbb{E}(Y \mid \mathcal{I}_1) \mid \mathcal{I}_2] = \mathbb{E}(Y \mid \mathcal{I}_2)$

(B) $\mathbb{E}(Y \mid \mathcal{I}_1) = \mathbb{E}(Y \mid \mathcal{I}_2)$ whenever $\mathcal{I}_1 \subset \mathcal{I}_2$

(C) $\mathbb{E}[\mathbb{E}(Y \mid \mathcal{I}_2)] = \mathbb{E}(Y)$

(D) $\mathbb{E}[\mathbb{E}(Y \mid \mathcal{I}_1)] = \mathbb{E}(Y \mid \mathcal{I}_2)$

5. Let (X, Y) be continuous random variables with the joint probability density function (PDF) defined as:

$$f_{X,Y}(x, y) = \begin{cases} 8xy, & 0 < y < x < 1 \\ 0, & \text{otherwise} \end{cases}$$

Which of the following statements correctly describes the independence of X and Y and the conditional expectation $E[Y|X = x]$?

(A) X and Y are independent, and $E[Y|X = x] = E[Y]$.

(B) X and Y are independent, and $E[Y|X = x] = \frac{2}{3}$.

(C) X and Y are not independent, and $E[Y|X = x] = \frac{1}{2}x$.

(D) X and Y are not independent, and $E[Y|X = x] = \frac{2}{3}x$.

6. Suppose (X, Y) follows a bivariate normal distribution and $\text{Cov}(X, Y) = 0$. Which statement is correct?

(A) X and Y are independent

(B) X and Y are uncorrelated but not independent

(C) The marginals must be uniform on $(0, 1)$

(D) $E(X \mid Y)$ is necessarily nonlinear in Y

7. Let $Z \sim N(0,1)$ be a standard normal random variable. Define

$$Z^+ = Z \mid (Z > 0),$$

that is, Z conditional on being positive.

Which of the following values is closest to $\mathbb{E}[Z^+]$?

(A) 0.40
(B) 0.60
(C) 0.80
(D) 1.00

8. Let X be a strictly positive random payoff with $\mathbb{E}[X] = 1$ and $\mathrm{Var}(X) > 0$. An agent has utility function $u(c) = \log c$.

The *certainty equivalent* c^{CE} is defined implicitly by

$$u(c^{CE}) = \mathbb{E}[u(X)].$$

Which statement is correct?

(A) $c^{CE} = \mathbb{E}[X]$ because expected utility equals utility of the mean.
(B) $c^{CE} > \mathbb{E}[X]$ because the agent prefers certainty.
(C) $c^{CE} < \mathbb{E}[X]$ because $\log(\cdot)$ is concave.
(D) $c^{CE} = 1$ for any random variable with mean one.

9. Let $X \sim N(0,1)$ and define the transformed random variable

$$Y = \begin{cases} X, & X \le 0, \\ 2X, & X > 0. \end{cases}$$

Let $\Phi(\cdot)$ denote the standard normal cumulative distribution function. Which of the following expressions correctly represents the cumulative distribution function $F_Y(y) = P(Y \le y)$ for all $y \in \mathbb{R}$?

(A) $F_Y(y) = \Phi(y)$ for all y

(B)

$$F_Y(y) = \begin{cases} \Phi(y), & y \le 0, \\ \Phi(y/2), & y > 0. \end{cases}$$

(C)

$$F_Y(y) = \begin{cases} \Phi(y), & y \le 0, \\ 2\Phi(y), & y > 0. \end{cases}$$

(D)

$$F_Y(y) = \begin{cases} \Phi(y), & y \le 0, \\ \Phi(2y), & y > 0. \end{cases}$$

10. Let $Z \sim N(\delta, 1)$ with $\delta = 3$ and define $W = Z^2$. Then W follows a noncentral chi-square distribution with 1 degree of freedom and noncentrality parameter $\lambda = \delta^2$. Which value equals $\mathbb{E}[W]$?

(A) 1

(B) 3

(C) 9

(D) 10

11. Suppose $\{X_1, X_2, \ldots, X_n\}$ is a sequence of i.i.d. random variables with $E[X_i] = \mu$ and $Var(X_i) = \sigma^2 < \infty$. Let $\overline{X}_n$ be the sample mean. Which of the following best describes the difference between the Law of Large Numbers (LLN) and the Central Limit Theorem (CLT)?

(A) The LLN states that $\overline{X}_n$ becomes a normal distribution as $n \to \infty$, while the CLT states that $\overline{X}_n$ converges to μ.

(B) The LLN describes the point to which $\overline{X}_n$ converges (the mean), while the CLT describes the shape of the distribution of $\overline{X}_n$ around that point as n increases.

(C) The LLN only applies to normal distributions, whereas the CLT applies to any distribution with finite variance.

(D) The LLN describes the behavior of the sample variance, while the CLT describes the behavior of the sample mean.

12. Let X be any random variable with $\mathbb{E}[X^4] < \infty$ and let ε satisfy $\mathbb{E}[\varepsilon \mid X] = 0$. Define $Y = X^2 + \varepsilon$. Which statement must be true?

(A) $\text{Cov}(Y, X) = 0$

(B) $\mathbb{E}[\varepsilon] = \mathbb{E}[X]$

(C) $\text{Cov}(\varepsilon, X^2) = 0$

(D) $\mathbb{E}[Y \mid X] = \mathbb{E}[Y]$

13. A financial analyst is testing the *weak-form efficient market hypothesis* by checking for serial correlation in daily stock returns ($H_0 : \rho = 0$ vs. $H_1 : \rho \neq 0$). The analyst is concerned that their test lacks the statistical power to detect small but profitable patterns in the data. Which of the following changes to the research design would successfully increase the power of the test to detect a specific non-zero correlation ρ_1?

(A) Changing the significance level α from 0.05 to 0.01 to ensure the discovered patterns are highly significant.

(B) Using a conservative non-parametric test instead of a parametric t-test to avoid normality assumptions.

(C) Extending the data set from 2 years of daily returns to 10 years of daily returns.

(D) Testing the hypothesis $H_0 : \rho = 0$ against a specific one-sided alternative $H_1 : \rho > 0$ when the true correlation is actually negative.

14. Suppose X has mean 0 and variance 1. Which statement best reflects the implication of Chebyshev's inequality?

(A) There exists a distribution of X such that $P(|X| \geq 2) = \frac{1}{4}$.

(B) For every distribution of X, $P(|X| \geq 2) < \frac{1}{4}$.

(C) Chebyshev's inequality implies $P(|X| \leq 2) = 0.95$.

(D) The inequality applies only when X is symmetric about zero.

15. Let $Y_1, Y_2, \ldots, Y_n$ be i.i.d. random variables with mean μ and variance σ^2. Define the sample mean

$$\overline{Y}_n = \frac{1}{n}\sum_{i=1}^{n} Y_i.$$

Consider the estimator

$$\widetilde{\mu} = \frac{Y_1 + Y_2}{2}.$$

Which statement is correct?

(A) $\widetilde{\mu}$ is biased and inconsistent

(B) $\widetilde{\mu}$ is unbiased but inefficient relative to $\overline{Y}_n$

(C) $\widetilde{\mu}$ is biased but has smaller variance than $\overline{Y}_n$

(D) $\widetilde{\mu}$ is unbiased and asymptotically efficient

16. Let Y_1, Y_2, Y_3 be independent and identically distributed random variables with mean μ and variance σ^2. Consider an unbiased linear estimator of μ of the form

$$\widehat{\mu} = aY_1 + bY_2 + cY_3,$$

where the constants a, b, c satisfy $a + b + c = 1$. Which choice of (a, b, c) minimizes $\mathrm{Var}(\widehat{\mu})$?

(A) $a = b = c = \frac{1}{3}$

(B) $a = \frac{1}{2},\ b = \frac{1}{2},\ c = 0$

(C) $a = 1,\ b = c = 0$

(D) Any triple satisfying $a + b + c = 1$

17. Suppose household food expenditures are normally distributed with mean \$20,000 and standard deviation \$500. Which probability is the largest?

(A) $P(X > 20{,}500)$

(B) $P(X < 19{,}500)$

(C) $P(|X - 20{,}000| > 500)$

(D) $P(|X - 20{,}000| < 500)$

18. A hedge fund manager is evaluating a potential *quantitative trading strategy* designed to generate "alpha" (excess returns). The manager tests the strategy against the null hypothesis that the strategy has *zero predictive power* (H_0 : Alpha = 0). In this financial context, a *Type II error* corresponds to:

    (A) Investing in the strategy because the data suggests it is profitable, when in reality it is just lucky and has no true predictive power.

    (B) Deciding not to invest in the strategy because the test results were not statistically significant, even though the strategy actually possesses a genuine edge.

    (C) Rejecting the null hypothesis at the 5% significance level when the p-value is 0.06.

    (D) Accurately identifying that a strategy is a "random walk" and choosing to use a low-cost index fund instead.

19. Let $X_1, X_2, \ldots, X_n$ be an i.i.d. sample from a population with a finite k-th moment. Which of the following statements correctly describes the properties of the Method of Moments (MoM) estimator $\widehat{\theta}_{MoM}$?

    (A) Under very general regularity conditions, MoM estimators are consistent, meaning $\widehat{\theta}_{MoM}$ converges in probability to the true parameter θ as $n \to \infty$.

    (B) MoM estimators are guaranteed to be the Best Linear Unbiased Estimators (BLUE) for any distribution within the exponential family.

    (C) A MoM estimator is always efficient, achieving the Cramér-Rao Lower Bound (CRLB) in large samples.

    (D) If a parameter θ represents the population variance, the MoM estimator $\widehat{\theta}_{MoM}$ will always be the unbiased sample variance $S^2 = \frac{1}{n-1}\sum(X_i - \overline{X})^2$.

20. An applied political economist analyzes a nationally representative survey estimating the vote share of Party A in an upcoming election. Based on a random sample, the estimated support for Party A is 48%. A 95% confidence interval for the true population support is reported as [43%, 53%].

    Which statement provides the *correct frequentist interpretation* of this confidence interval?

    (A) The true population support for Party A is exactly between 43% and 53%.

    (B) There is a 95% probability that the true support lies in the interval [43%, 53%].

    (C) Once calculated, the interval [43%, 53%] is fixed; the true parameter is either in it or it is not, so the probability is 1 or 0.

    (D) The survey provides statistical evidence that Party A is unpopular.

Part II. Analytical Problems (40 points)

Instructions: Detailed calculations or justifications must be clearly written for each item; otherwise, no points will be awarded. Answers must be written in the provided **answer sheet**, marked with the corresponding problem number and item label. Please answer the questions sequentially and **do not** write your answers on this question paper.

Problem 1. (20 points)

Let $X \sim \text{Unif}(0,1)$ with density $f_X(x) = 1$ for $0 < x < 1$. Conditional on $X = x$, the random variable Y satisfies

$$\mathbb{E}[Y \mid X = x] = x, \qquad \text{Var}(Y \mid X = x) = x.$$

Answer the following questions:

(a) (5 points) Using only the provided conditional moments, *calculate the numerical value* of $\mathbb{E}[Y]$. Then, *verify numerically* whether Y and X are uncorrelated by computing their covariance. Clearly state any statistical identities and marginal moments of X used in your derivation.

(b) (5 points) *Compute the numerical value* of $\text{Var}(Y)$ using the law of total variance. Then explain which component of the variance decomposition captures *heterogeneity driven by X* and which captures *idiosyncratic risk.*

(c) (6 points) Using the information provided about the conditional moments of Y given X, *compute the numerical value* of the unconditional second moment $\mathbb{E}[Y^2]$.

(i) First, obtain $\mathbb{E}[Y^2]$ by combining your previously computed values of $\mathbb{E}[Y]$ and $\text{Var}(Y)$.

(ii) Next, compute $\mathbb{E}[Y^2]$ by evaluating the conditional second moment $\mathbb{E}[Y^2 \mid X]$ and then taking expectations with respect to X.

Verify that the two calculations give the same numerical value, and explain why this agreement is not coincidental.

(d) (4 points) Discuss whether the joint distribution of (X, Y) is uniquely identified by the information given in the problem. Provide a brief explanation of what additional assumptions would be required to fully characterize the conditional distribution of $Y \mid X$.

Problem 2. (20 points)

Let $X_1, X_2, \ldots, X_n$ be an i.i.d. sample from a Bernoulli distribution with

$$P(X_i = 1 \mid p) = p, \qquad 0 < p < 1,$$

and let $\theta = p$ denote the unknown parameter. Assume all regularity conditions required for the Cramér–Rao inequality hold.

Answer the following questions:

(a) (8 points) Derive the log-likelihood function and the score function for θ. Verify explicitly that the regularity condition

$$E_\theta\left[\frac{\partial}{\partial\theta} \log f(X_i \mid \theta)\right] = 0$$

holds for the Bernoulli model.

(b) (8 points) Compute the Fisher Information for a single observation and for the full sample *using both* of the following definitions:

$$I(\theta) = \mathrm{Var}_\theta\left(\frac{\partial}{\partial\theta}\log f(X_i \mid \theta)\right), \qquad I(\theta) = -E_\theta\left(\frac{\partial^2}{\partial\theta^2}\log f(X_i \mid \theta)\right).$$

Show that the two expressions coincide.

(c) (4 points) State the Cramér–Rao Lower Bound for any unbiased estimator of p. Provide the explicit expression of the bound as a function of p and n, and state the necessary condition under which an unbiased estimator attains the CRLB.

— [End of the Exam Question Paper] —

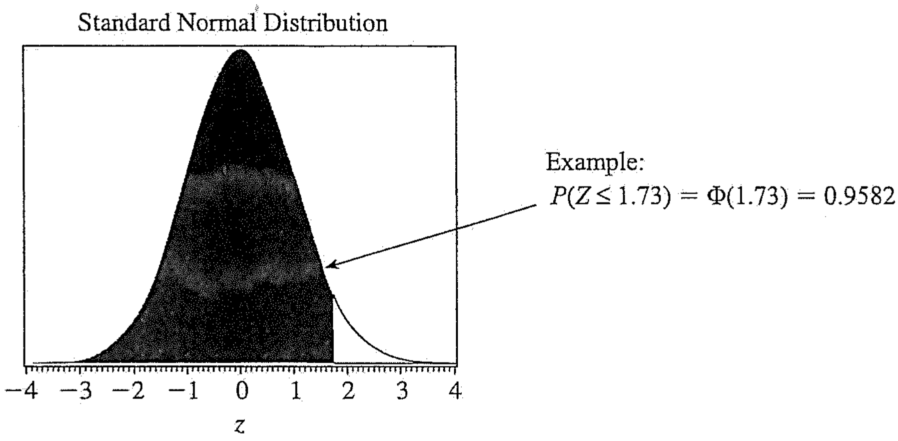

Table 1 **Cumulative Probabilities for the Standard Normal Distribution**
$\Phi(z) = P(Z \leq z)$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990