Instructions:

- Write in Chinese or English only.
- Make sure all your answers are legible, unquestionably labeled, and clearly explained.
- A standard normal probability table is attached on the last page.

1. **(10 points)** A fast-food chain uses online and telephone questionnaires to estimate average customer satisfaction at one of its stores. Each customer reports an overall satisfaction rating between 1 and 10. The fast-food chain will solicit 66 online surveys and 42 telephone surveys. The results of each survey are independent draws from the same distribution. Let $\overline{X}$ and $\overline{Y}$ denote the sample mean responses to the online and telephone surveys, respectively. The fast-food chain would like to take a weighted average $\overline{C} = w\overline{X} + (1-w)\overline{Y}$ of these sample means to obtain a single estimator of average customer satisfaction. Determine the choice of weight w that leads to the most efficient unbiased estimator of average customer satisfaction.

2. **(20 points)** A bank manager is analyzing the usage patterns of her bank's ATM. She observes that the amount of time (X, in minutes) a user spends at the ATM follows the probability distribution below:

x	$\Pr(X = x)$
2	0.5
3	0.2
4	0.3

The amounts of time spent by different users at the ATM are independent of one another. Calculate the probability that the total time spent by the first 120 users at the ATM is at least 330 minutes.

3. **(20 points)** A power company claims that it has successfully reduced the average carbon dioxide emissions from one of its facilities to 24,000 metric tons per day to meet its obligations in an emissions trading program. However, an emissions regulator suspects that the actual average is lower, at 23,000 metric tons per day. Assume the regulator's estimate is correct, and that daily emissions are independent and identically distributed with a standard deviation of 2,200 metric tons. Determine the number of days required to perform a one-sided hypothesis test with a Type I error probability of 0.01 and a Type II error probability of 0.02.

4. A researcher is analyzing the factors influencing whether individuals participate in a job training program. The binary variable y_i equals 1 if individual i participates and 0 otherwise. The researcher uses the following probit model:

$$\mathbb{P}(y_i = 1|x_i) = \Phi(x_i^\top \beta), \tag{1}$$

where

- $\Phi(\cdot)$ is the cumulative distribution function (CDF) of the standard normal distribution,
- $x_i = [x_{i1}, x_{i2}, \ldots, x_{ik}]^\top$ is a $k \times 1$ vector of explanatory variables (e.g., age, education, and income), and
- $\beta = [\beta_1, \beta_2, \ldots, \beta_k]^\top$ is a vector of coefficients to be estimated.

(a) **(8 points)** Derive the log-likelihood function and the corresponding first-order conditions for the probit model given a sample of n observations.

(b) **(4 points)** Derive the expression for the marginal effect of a continuous explanatory variable x_{ik} on the probability $\mathbb{P}(y_i = 1|x_i)$.

5. A researcher is studying the relationship between years of education (Edu_i) and earnings (Earn_i) using the following linear regression model:

$$\text{Earn}_i = \beta_0 + \beta_1 \text{Edu}_i + \beta_2 \text{Ability}_i + \epsilon_i,\ i = 1, 2, \ldots, n, \tag{2}$$

in which Ability_i is positively correlated with both Edu_i and Earn_i, and the error term $\epsilon_i \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma^2)$ is uncorrelated with both Edu_i and Ability_i. However, the researcher does not observe Ability_i.

(a) **(6 points)** Suppose that the researcher considers a simplified model:

$$\text{Earn}_i = b_0 + b_1 \text{Edu}_i + \varepsilon_i,\ i = 1, 2, \ldots, n. \tag{3}$$

Will the ordinary least squares (OLS) estimate of b_1 in the regression above be a consistent estimator for β_1? If not, in which direction does the bias occur? Explain your answer with relevant equations.

(b) **(6 points)** Suppose that the researcher has access to two measured variables, $\widehat{\text{Ability}}_i$ and $\widetilde{\text{Ability}}_i$:

$$\widehat{\text{Ability}}_i = \text{Ability}_i + u_i, \text{ and } \widetilde{\text{Ability}}_i = \text{Ability}_i + v_i, \tag{4}$$

where error terms $u_i \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma_u^2)$ and $v_i \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma_v^2)$ are uncorrelated with Edu_i, Ability_i, ϵ_i, and each other. To minimize the effects of these errors, the researcher calculates the average of these two measures:

$$\overline{\text{Ability}}_i = \frac{\widehat{\text{Ability}}_i + \widetilde{\text{Ability}}_i}{2}, \tag{5}$$

and considers the following model:

$$\text{Earn}_i = \gamma_0 + \gamma_1 \text{Edu}_i + \gamma_2 \overline{\text{Ability}}_i + e_i,\ i = 1, 2, \ldots, n. \tag{6}$$

Will the ordinary least squares (OLS) estimate of γ_1 in the regression above be a consistent estimator for β_1? Will the OLS estimate of γ_2 be a consistent estimator for β_2? Explain your answer with relevant equations.

(c) **(6 points)** Besides the estimators above, how would you consistently estimate β_1 and β_2? Provide a detailed explanation supported by relevant equations and formulas.

接次頁

6. Consider the following dynamic panel data model:

$$y_{it} = \mu_i + \alpha y_{i,t-1} + \epsilon_{it},\ i = 1, 2, \ldots, n, \text{ and } t = 2, 3, \ldots, T, \tag{7}$$

where $T \geq 4$,

- y_{it} is the dependent variable for individual i at time t,
- $y_{i,t-1}$ is the lagged dependent variable,
- μ_i represents the individual-specific effect, and
- ϵ_{it} is the idiosyncratic error term, assumed to be i.i.d. $\mathcal{N}(0, \sigma^2)$.

Taking the first difference of equation (7) eliminates the individual-specific effect, resulting in:

$$\Delta y_{it} = \alpha \Delta y_{i,t-1} + \Delta \epsilon_{it},\ i = 1, 2, \ldots, n, \text{ and } t = 3, 4, \ldots, T, \tag{8}$$

where $\Delta y_{it} = y_{it} - y_{i,t-1}$ represents the change in y over time. The first-difference (FD) estimator of α, denoted as $\hat{\alpha}_{FD}$, is obtained by applying ordinary least squares (OLS) to the transformed variables Δy_{it} and $\Delta y_{i,t-1}$:

$$\hat{\alpha}_{FD} = \frac{\sum_{i=1}^{n} \sum_{t=3}^{T} \Delta y_{i,t-1} \Delta y_{it}}{\sum_{i=1}^{n} \sum_{t=3}^{T} \Delta y_{i,t-1}^2}. \tag{9}$$

(a) **(8 points)** (Nickell, 1981, *Econometrica*) Consider the scenario where $n \to \infty$ while T remains fixed. Explain whether $\hat{\alpha}_{FD}$ is consistent. Justify your answer.

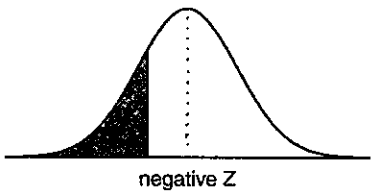
(b) **(4 points)** Now consider the scenario where both $n \to \infty$ and $T \to \infty$. Determine whether $\hat{\alpha}_{FD}$ is a consistent estimator, and explain your reasoning.

(c) **(8 points)** (Anderson and Hsiao, 1982, *Journal of Econometrics*) The Anderson and Hsiao (AH) estimator of α, denoted as $\hat{\alpha}_{AH}$, is obtained by applying two stage least squares (2SLS) to the transformed variables Δy_{it} and $\Delta y_{i,t-1}$, using $\Delta y_{i,t-2}$ as the instrumental variable (IV):

$$\hat{\alpha}_{AH} = \frac{\sum_{i=1}^{n} \sum_{t=4}^{T} \Delta y_{i,t-2} \Delta y_{it}}{\sum_{i=1}^{n} \sum_{t=4}^{T} \Delta y_{i,t-2} \Delta y_{i,t-1}}. \tag{10}$$

Again, consider the scenario where $n \to \infty$ while T remains fixed. Discuss whether $\hat{\alpha}_{AH}$ is a consistent estimator for α, and provide justification for your conclusion.

Standard normal probability table



Second decimal place of Z										
0.09	0.08	0.07	0.06	0.05	0.04	0.03	0.02	0.01	0.00	Z
0.0002	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	−3.4
0.0003	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0005	0.0005	0.0005	−3.3
0.0005	0.0005	0.0005	0.0006	0.0006	0.0006	0.0006	0.0006	0.0007	0.0007	−3.2
0.0007	0.0007	0.0008	0.0008	0.0008	0.0008	0.0009	0.0009	0.0009	0.0010	−3.1
0.0010	0.0010	0.0011	0.0011	0.0011	0.0012	0.0012	0.0013	0.0013	0.0013	−3.0
0.0014	0.0014	0.0015	0.0015	0.0016	0.0016	0.0017	0.0018	0.0018	0.0019	−2.9
0.0019	0.0020	0.0021	0.0021	0.0022	0.0023	0.0023	0.0024	0.0025	0.0026	−2.8
0.0026	0.0027	0.0028	0.0029	0.0030	0.0031	0.0032	0.0033	0.0034	0.0035	−2.7
0.0036	0.0037	0.0038	0.0039	0.0040	0.0041	0.0043	0.0044	0.0045	0.0047	−2.6
0.0048	0.0049	0.0051	0.0052	0.0054	0.0055	0.0057	0.0059	0.0060	0.0062	−2.5
0.0064	0.0066	0.0068	0.0069	0.0071	0.0073	0.0075	0.0078	0.0080	0.0082	−2.4
0.0084	0.0087	0.0089	0.0091	0.0094	0.0096	0.0099	0.0102	0.0104	0.0107	−2.3
0.0110	0.0113	0.0116	0.0119	0.0122	0.0125	0.0129	0.0132	0.0136	0.0139	−2.2
0.0143	0.0146	0.0150	0.0154	0.0158	0.0162	0.0166	0.0170	0.0174	0.0179	−2.1
0.0183	0.0188	0.0192	0.0197	0.0202	0.0207	0.0212	0.0217	0.0222	0.0228	−2.0
0.0233	0.0239	0.0244	0.0250	0.0256	0.0262	0.0268	0.0274	0.0281	0.0287	−1.9
0.0294	0.0301	0.0307	0.0314	0.0322	0.0329	0.0336	0.0344	0.0351	0.0359	−1.8
0.0367	0.0375	0.0384	0.0392	0.0401	0.0409	0.0418	0.0427	0.0436	0.0446	−1.7
0.0455	0.0465	0.0475	0.0485	0.0495	0.0505	0.0516	0.0526	0.0537	0.0548	−1.6
0.0559	0.0571	0.0582	0.0594	0.0606	0.0618	0.0630	0.0643	0.0655	0.0668	−1.5
0.0681	0.0694	0.0708	0.0721	0.0735	0.0749	0.0764	0.0778	0.0793	0.0808	−1.4
0.0823	0.0838	0.0853	0.0869	0.0885	0.0901	0.0918	0.0934	0.0951	0.0968	−1.3
0.0985	0.1003	0.1020	0.1038	0.1056	0.1075	0.1093	0.1112	0.1131	0.1151	−1.2
0.1170	0.1190	0.1210	0.1230	0.1251	0.1271	0.1292	0.1314	0.1335	0.1357	−1.1
0.1379	0.1401	0.1423	0.1446	0.1469	0.1492	0.1515	0.1539	0.1562	0.1587	−1.0
0.1611	0.1635	0.1660	0.1685	0.1711	0.1736	0.1762	0.1788	0.1814	0.1841	−0.9
0.1867	0.1894	0.1922	0.1949	0.1977	0.2005	0.2033	0.2061	0.2090	0.2119	−0.8
0.2148	0.2177	0.2206	0.2236	0.2266	0.2296	0.2327	0.2358	0.2389	0.2420	−0.7
0.2451	0.2483	0.2514	0.2546	0.2578	0.2611	0.2643	0.2676	0.2709	0.2743	−0.6
0.2776	0.2810	0.2843	0.2877	0.2912	0.2946	0.2981	0.3015	0.3050	0.3085	−0.5
0.3121	0.3156	0.3192	0.3228	0.3264	0.3300	0.3336	0.3372	0.3409	0.3446	−0.4
0.3483	0.3520	0.3557	0.3594	0.3632	0.3669	0.3707	0.3745	0.3783	0.3821	−0.3
0.3859	0.3897	0.3936	0.3974	0.4013	0.4052	0.4090	0.4129	0.4168	0.4207	−0.2
0.4247	0.4286	0.4325	0.4364	0.4404	0.4443	0.4483	0.4522	0.4562	0.4602	−0.1
0.4641	0.4681	0.4721	0.4761	0.4801	0.4840	0.4880	0.4920	0.4960	0.5000	−0.0

*For $Z \leq -3.50$, the probability is less than or equal to 0.0002.

試題隨卷繳回