Econometrics
Midterm Exam, 2020
1 Hour and 45 Minutes

Critical Values		
$F_{2,N,1-\alpha} = 3.00$	$F_{3,N,1-\alpha} = 2.60$	$F_{6,N,1-\alpha} = 2.10$
$t_{12,1-\alpha/2} = 2.18$	$t_{27,1-\alpha/2} = 2.05$	$t_{21,1-\alpha/2} = 2.08$
$\alpha = 0.05$, $N > 200$		

Question 1. (15 points) Consider the following model: $Y = \alpha + \beta X + U$, where $E(U|X) = 0$, $E\left(U^2|X\right) = \sigma^2 > 0$ and the conditional distribution of U given X is $N\left(0, \sigma^2\right)$.

(a) (5 points) What is $E(Y|X)$? What is the conditional distribution of Y given X?

(b) (5 points) In this question, use the following result: if the conditional distribution of Y given X is $N\left(m(X), s(X)^2\right)$ $(s(X) > 0)$, then the conditional distribution of $\frac{Y-m(X)}{s(X)}$ given X is $N(0,1)$ and

$$P\left(\frac{Y - m(X)}{s(X)} \leq z|X\right) = \Phi(z),$$

where $\Phi(z)$ is the standard normal CDF. Here, the left hand side means the conditional probability of $\frac{Y-m(X)}{s(X)} \leq z$ given X. For a random variable Z, its τ-th quantile $(\tau \in (0,1))$ q_τ is defined by the equation: $P(Z \leq q_\tau) = \tau$. Similarly, a function $q_\tau(X)$ is the τ-th quantile of the conditional distribution of Y given X if

$$P(Y \leq q_\tau(X)|X) = \tau.$$

Find the expression of $q_\tau(X)$. Hint: Let z_τ denote the τ-th quantile of the standard normal distribution so that $\Phi(z_\tau) = \tau$.

(c) (5 points) Suppose that $E\left(U^2|X\right) = e^{2X}$ and the conditional distribution of U given X is $N\left(0, e^{2X}\right)$. Find the expression of $q_\tau(X)$.

Question 2. (25 points) Consider the following regression model

$$\begin{aligned} Y_i &= \beta X_i + U_i, \ i = 1, \ldots, n; \\ E(U_i|X_1, \ldots, X_n) &= 0; \\ E\left(U_i^2|X_1, \ldots, X_n\right) &= \sigma^2; \\ E(U_i U_j|X_1, \ldots, X_n) &= 0 \text{ for all } i \neq j. \end{aligned}$$

Assume that $\bar{X} \neq 0$ and consider the following estimator of β:

$$\tilde{\beta} = \frac{\bar{Y}}{\bar{X}}, \text{ where } \bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i \text{ and } \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

(a) (10 points) Show that $\tilde{\beta}$ is unbiased.

(b) (10 points) Show that

$$Var\left(\tilde{\beta}|X_1, \dots, X_n\right) = \frac{\sigma^2}{n\left(\bar{X}\right)^2}.$$

(c) (5 points) Let

$$\hat{\beta} = \frac{\sum_{i=1}^{n} X_i Y_i}{\sum_{i=1}^{n} X_i^2},$$

and recall that

$$Var\left(\hat{\beta}|X_1, \dots, X_n\right) = \frac{\sigma^2}{\sum_{i=1}^{n} X_i^2}.$$

Without using the Gauss-Markov Theorem, show that

$$Var\left(\hat{\beta}|X_1, \dots, X_n\right) \leq Var\left(\tilde{\beta}|X_1, \dots, X_n\right).$$

Hint: What is the relationship between $\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2$, $\sum_{i=1}^{n} X_i^2$, and $n\left(\bar{X}\right)^2$?

Question 3. (15 points) Find A-C in the Stata output below. The number of observations is 29. Justify your answer. According to the results, is x significant? The significance level is 5%. Explain.

y	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
x	A	.373891	B	0.020	.2025939	C
_cons	1.261567	.1814382	6.95	0.000	.8892861	1.633847

Question 4. (10 points) Consider a simple classical linear regression model:

$$\begin{aligned}
\log(Y_i) &= \beta_0 + \beta_1 \log(X_i) + U_i, \; i = 1, \dots, n, \\
E(U_i|X_1, \dots, X_n) &= 0, \\
E(U_i^2|X_1, \dots, X_n) &= \sigma^2, \\
E(U_i U_j|X_1, \dots, X_n) &= 0 \text{ for } i \neq j.
\end{aligned}$$

Let $\hat{\beta}_1$ denote the OLS estimator obtained from a regression of $\log(Y_i)$ against $\log(X_i)$. For some known constant $k > 0$, let $\tilde{\beta}_1$ denote the OLS estimator obtained from a regression of $\log(Y_i)$ against $\log(kX_i)$. Find the relationship between $Var(\hat{\beta}_1|X_1, \dots, X_n)$ and $Var(\tilde{\beta}_1|X_1, \dots, X_n)$.

Question 5. (25 Points) We are interested in explaining a worker's wage in terms of the number of years of education (educ) and years of experience (exper) using the following model:

$$log(wage) = \beta_0 + \beta_1 educ + \beta_2 exper + u.$$

The estimated parameters by OLS for a sample of $n = 935$ observations are displayed in the table.

Several extensions of this model were considered to address the effects of being married (with the binary variable married) and/or being black (with binary black) or possible nonlinearity on the effect of years of experience.

(a) Test whether the wage regressions for married workers and unmarried workers are the same. Hint: Perform a F test. Model 2 is the unrestricted and Model 1 is the restricted.

(b) Based on a statistical test, do the effects of education and experience depend on the marriage status? Hint: Perform a F test. Model 2 is the unrestricted and Model 3 is the restricted.

(c) What we conclude about the possible nonlinearity of the relationship of $log(wage)$ with respect to the years of experience? Can you conclude that years of experience has no significant effect on $log(wage)$ in Model 5? Make two statistical tests to answer these questions. Hint: Look at Model 5 and Model 6. Use a t test to answer the first question and use a F test to answer the second question.

Variables	*Model 1*	*Model 2*	*Model 3*	*Model 5*	*Model 6*
educ	*0.07778*	*0.05316*	*0.07815*	*0.071984*	*0.05571*
	(0.00669)	*(0.02085)*	*(0.00653)*	*(0.00677)*	*(0.00600)*
exper	*0.01977*	*0.00038*	*0.01829*	*0.01678*	
	(0.00330)	*(0.01066)*	*(0.00330)*	*(0.01389)*	
educ×married		*0.02813*			
		(0.02194)			
exper×married		*0.01952*			
		(0.01120)			
married		*-0.38069*	*0.20926*	*0.18873*	*0.21311*
		(0.36818)	*(0.04272)*	*(0.04763)*	*(0.04709)*
black				*-0.24128*	*-0.22500*
				(0.08417)	*(0.08212)*
married×black				*0.03543*	*0.01071*
				(0.09404)	*(0.09224)*
exper^2				*0.0000486*	
				(0.00058)	
constant	*5.50271*	*5.85694*	*5.32796*	*5.46653*	*5.86609*
	(0.11427)	*(0.34889)*	*(0.11574)*	*(0.12914)*	*(0.09445)*
observations	*935*	*935*	*935*	*935*	*935*
R^2	*0.13086*	*0.15705*	*0.15420*	*0.18132*	*0.15417*

Question 6. (10 points) Suppose that we observe a random sample $X_1, ..., X_n$ where $X_1, ..., X_n$ are independent $N\left(\mu, \sigma^2\right)$ random variables. σ^2 is known. Consider the test statistic ($\overline{X} = n^{-1}\sum_{i=1}^{n} X_i$)

$$T = \frac{\overline{X}}{\sigma/\sqrt{n}}$$

for the two-sided hypothesis test $H_0 : \mu = 0$ against $H_1 : \mu \neq 0$. H_0 is rejected if $|T| > z_{1-\alpha/2}$ where $z_{1-\alpha/2}$ is the $(1-\alpha/2)$-th quantile of $N(0, 1)$. Give the expression of the power function as a function of μ. Use Φ to denote the standard normal CDF.