Problem 1 (21 points)

An economist runs the regression

$$y_i = \beta_0 + x_{1i}\beta_1 + x_{2i}\beta_2 + x_{3i}\beta_3 + \varepsilon_i \tag{1}$$

The results are summarized in the following table:

Equation 1.		
Variable	Coefficient	Std. Error
β_0	0.82	1.04
β_1	-1.16	5.74
β_2	7.63	3.70
β_3	-2.57	0.94
R-squared	0.102	
Sum squared resid	10297	
Number of observations	100	

(a) (3 points) Construct a 95% confidence interval for β_2 in Equation 1.

(b) (3 points) Test whether β_1 in Equation 1 equals 0 (against the alternative that it differs from 0). Test at a 5% level of significance.

(c) (6 points) How would you test whether $\beta_1 = \beta_2 = \beta_3 = 0$ in equation 1? How does your answer depend on the assumptions made on ε_i?

.

Using the same data, the economist also runs the regression

$$y_i = \beta_0 + x_{2i}\beta_2 + \nu_i \tag{2}$$

Some of the results are summarized in the following table:

Equation 2.		
Variable	Coefficient	Std. Error
β_0	0.83	1.07
β_2	6.86	3.81
R-squared	?	
Sum squared resid	11095	
Number of observations	100	

(d) (9 points) What is R^2 in equation 2?

Problem 2 (14 points)

Suppose that the true relationship between four variables y, x, z and w is given by

$$
\begin{aligned}
y_i &= \delta_0 + \delta_1 z_i + \delta_2 x_i + u_i \qquad (3) \\
z_i &= \alpha_0 + \alpha_1 x_i + \alpha_2 w_i + v_i \qquad (4)
\end{aligned}
$$

where the unobserved error terms u_i and v_i both have mean 0 and are uncorrelated with x_i and w_i. It is not assumed that u_i and v_i are uncorrelated. Answer the questions below under the assumption of random sampling of (y_i, x_i, z_i, w_i) (i.e., the observations are *i.i.d.*).

(a) (7 points) Suppose that you regress y_i on a constant, z_i and x_i. Are the regression coefficient on z_i and x_i consistent estimators of δ_1 and δ_2? Explain. If not, how would you estimate δ_1 and δ_2? (State exactly what additional assumptions, if any, you need to make)

(b) (7 points) Assume that w_i and x_i are independent, and consider the regression

$$y_i = \beta_0 + \beta_1 x_i + \text{“error”}$$

Interpret the coefficients β_0 and β_1 in terms of the coefficients in (3) and (4).

Problem 3 (15 points)

Consider the model

$$y_i = z_i\delta + \varepsilon_i, \qquad \text{with} \qquad E\left[\varepsilon_i z_i\right] = 0$$

Moreover assume that there is another variable x_i such that $E\left[\varepsilon_i x_i\right] = 0$. Both z_i and x_i are one–dimensional. Assume that assumptions (3.1)–(3.5) of Hayashi are satisfied and that

$$E\left[\begin{pmatrix} x_i \\ z_i \end{pmatrix}\begin{pmatrix} x_i \\ z_i \end{pmatrix}'\right] = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$$

and

$$E\left[\varepsilon_i^2 \middle| x_i, z_i\right] = 1$$

(a) (5 points) Find the asymptotic distribution of the OLS estimator in a regression of y_i on z_i.

(b) (5 points) Find the asymptotic distribution of the 2SLS estimator of δ that uses x_i as instrument.

(c) (5 points) Find the asymptotic distribution of the OLS estimator in a regression of y_i on x_i.

Problem 4 (10 points)

Consider two linear regression models

$$y_t = x_t'\beta + \varepsilon_t$$

and

$$w_t = z_t'\delta + \nu_t$$

Assume that the standard assumptions for asymptotics in regression models (Hayashi assumptions 2.1–2.5) are satisfied for each of the models.

Suppose that $\widetilde{\beta}$ and $\widetilde{\delta}$ are consistent estimators of β and δ (but they are not necessarily the OLS estimators). Is

$$T^{-1} \sum_{t=1}^{T} \left(y_t - x_t'\widetilde{\beta} \right) \left(w_t - z_t'\widetilde{\delta} \right)$$

necessarily a consistent estimator of the covariance between ε_t and ν_t? Explain.

Problem 5 (15 points)

Suppose $Y_i|\mu \sim iid\ N(\mu, 1),\ i = 1, ..., n,\ \mu \sim N(2, 1),$ and $L(\widehat{\mu}, \mu) = (\widehat{\mu} - \mu)^2$. Suppose $n = 1$.

(a) (3 points) What is the MLE of μ? (You do not need to derive the answer; you may just state it.)

(b) (3 points) Show the Bayes risk of the MLE.(You do not need to derive the answer, just state it.)

(c) (3 points) Show the posterior distribution for μ. (You do not need to derive the answer; just may state it.)

(d) (3 points) Show the Bayes estimator of μ.(You do not need to derive the answer; you may just state it.)

(e) (3 points) Show the Bayes risk of the Bayes estimator. (You do not need to derive the answer; you may just state it.)

Problem 6 (5 points)

X, Y, U are independent random variables: $X \sim N(5, 4)$, $Y \sim \chi^2_1$ and U is distributed Bernoulli with $P(U = 1) = 0.3$. Let $W = UX + (1 - U)Y$. Find the mean and variance of W.

Problem 7 (10 points)

Y_i are *iid* Bernoulli random variables with $P(Y_i = 1) = p$. I am interested in $H_0 : p = 1/2$ versus $H_a : p > 1/2$. The sample size is $n = 100$, and I decide to the reject the null hypothesis if $\sum_{i=1}^{n} Y_i > 55$.

(a) (5 points) Use the central limit theorem to derive the approximate size of the test.

(b) (5 points) Suppose $p = 0.60$. Use the central limit theorem to derive the power of the test.

Problem 8 (30 points)

Suppose $Y_i|\mu \sim iid\ N(\mu, 1)$, $i = 1, ..., n$. Suppose that you know $\mu \leq 5$. Let $\widehat{\mu}_{MLE}$ denote the value of μ that maximizes the likelihood, subject to the constraint $\mu \leq 5$, and let $\overline{Y}$ denote the sample mean of the Y_i.

(a) (12 points) Show that $\widehat{\mu}_{MLE} = a(\overline{Y})5 + (1 - a(\overline{Y}))\overline{Y}$, where $a(\overline{Y}) = 1$ for $\overline{Y} \geq 5$, and $a(\overline{Y}) = 0$ otherwise. (Hint: you might find it useful to write $(y_i - \mu) = (y_i - \overline{Y}) + (\overline{Y} - \mu)$.)

(b) (14 points) Suppose that $\mu = 3$.

(b.i) (6 points) Show that $a(\overline{Y}) \xrightarrow{p} 0$. (Hint: What is $P(a(\overline{Y}) = 0)$?)

(b.ii) (4 points) Show that $\widehat{\mu}_{MLE} \xrightarrow{p} \mu$.

(b.iii) (4 points) Show that $\sqrt{n}(\widehat{\mu}_{MLE} - \mu) \xrightarrow{d} N(0,1)$.

(c) (4 points) Suppose that $\mu = 5$. Does $\widehat{\mu}_{MLE} \xrightarrow{p} \mu$? Explain