Studienzentrum Gerzensee Doctoral Program in Economics Econometrics
2020-21
Mark Watson

(This version: August 23, 2020)

This is a list of exercises that I have previously used as assignments and exams. You will see that there a lot of repetition.

1. Let $B_1, \ldots, B_m$ be mutually exclusive events each with positive probability, $B = \bigcup_{i=1}^{m} B_i$, and suppose $\mathbb{P}(A|B_i) = \mathbb{P}(A|B_1)$ for $i = 1, \ldots, m$. Show that $\mathbb{P}(A|B) = \mathbb{P}(A|B_1)$.

2. A and B are two events. Let $C = A \cup B$ and $D = A \cap B$.

   (a) Show $\mathbb{P}(C) \leq \mathbb{P}(A) + \mathbb{P}(B)$

   (b) Show $\mathbb{P}(D) \geq 1 - \mathbb{P}(A^c) - \mathbb{P}(B^c)$. (Hint: show $A \cap B = (A^c \cup B^c)^c$ and apply the result in (a).)

   (c) Let $A_1, A_2, \ldots A_n$ denote n events. Show $\mathbb{P}(A_1 \cup A_2 \cup \ldots \cup A_n) \leq \mathbb{P}(A_1) + \mathbb{P}(A_2) + \ldots + \mathbb{P}(A_n)$. (Hint: Use induction. You showed this result in (a) for $n = 2$. Show that if it holds for $n - 1$, then it holds for n.) This is sometimes called Bonferroni's inequality.

3. Suppose X has probability density $f(x) = \frac{1}{4}$ for $1 < x < 5$ and $f(x) = 0$ elsewhere. Let $Y = X^4$.

   (a) What is the CDF of Y?

   (b) What is the probability density function of Y?

4. Suppose X has probability density $f(x) = \frac{1}{4}$ for $-1 < x < 3$ and $f(x) = 0$ elsewhere. Let $Y = X^2$.

   (a) What is the CDF of Y?

   (b) What is the probability density function of Y?

   (c) What is the value of $\mathbb{E}(Y)$

5. The random variable X has the *inverse Gamma* distribution, with density $f_X(x) \propto x^{-(\alpha+1)} \exp\left(-\frac{\beta}{x}\right)$ for $x > 0$ and zero elsewhere, and where α and β are positive constants. Write this as $X \sim \mathcal{IG}(\alpha, \beta)$ The random variable Y has the *Gamma* distribution, with density $f_Y(y) = y^{a-1} \exp\left(-\frac{y}{b}\right)$ for $y > 0$ and zero elsewhere, and where a and b are positive constants. Write this as $Y \sim \mathcal{G}(a, b)$. Show that if $Y \sim \mathcal{G}(a, b)$ then $X = (1/Y) \sim \mathcal{IG}(a, b^{-1})$. (This result is useful in Bayes analysis of normal models, where the prior and posterior for σ^2 have an $\mathcal{IG}$ distribution. Draws of σ^2 can be generated using the reciprocals of draws from the $\mathcal{G}$ distribution.)

6. The joint density of X and Y is given by $f_{X,Y}(x, y) = c(x + y)$ for $0 < x < 1$ and $0 < y < 1$, and is equal to zero elsewhere.

   (a) Find the constant c.

   (b) Find the marginal density of X.

   (c) Find $\mathbb{P}(X > Y)$.

(d) Find the mean and variance of Y given $X = 0.3$.

7. X is a random variable with density $f(x) = (\frac{1}{2})^x$, $x = 1, 2, 3, ...$ and 0 elsewhere. Find $\mathbb{E}(X)$.

8. Suppose that X is a continuous random variable with CDF $F(x)$ and density$f(x)$. Show that $\mathbb{E}(X) = \int_0^1 F^{-1}(t)dt$.

9. Suppose $f_{X,Y}(x, y) = 1/4$ for $1<x<3$ and $1<y<3$, and $f_{X,Y}(x, y) = 0$ elsewhere.

   (a) Derive the marginal densities of X and Y.

   (b) Find $\mathbb{E}(Y|X = 2.3)$.

   (c) Find $P[(X<2.5)\cap(Y<2.0)]$.

   (d) Let $W = X + Y$. Derive the moment generating function of W.

10. $f_X(x) = 2/x^3$ for $x>1$ and zero elsewhere.

    (a) Compute the CDF of X.

    (b) (b) Find $\mathbb{E}(X)$.

    (c) Find $\mathbb{E}(1/X)$.

    (d) Let $Y = 1/X$.

        i. Find $f_Y(y)$.

        ii. Find $\mathbb{E}(Y)$ (and verify that the answer is the same as your answer to (c).)

11. X is a continuous random variable with density $f(x)$ and CDF $F(x)$ where F is $1 - to - 1$. Let $Y = F(X)$.

    (a) Show that $Y \sim \mathcal{U}[0, 1]$. (Y is called the *probability integral transform* (PIT) of X).

    (b) A forecasting agency publishes "density forecasts" (or "predictive densities") for a variable, say X_{t+h}, which they claim are accurate estimates of the density of X_{t+h} given information at time t. Call these densities $f_{t+h|t}(x)$. Suppose you have a history of these forecasts for $t = 1, \ldots, T$. Discuss how the result in (a) might be useful for evaluating the accuracy of these probability densities.

12. A survey is conducted in which a "Yes-or-No" question is asked about a sensitive topic: for example "Have you ever cheated on an exam?". Respondents may be reticent to answer Yes when that is the true answer, and this reticence results in biased surveys. To mitigate this problem researchers sometimes use "Random Response Questions". For these questions the respondent is instructed to flip a coin whose outcome is not observed by the surveyor, to answer Yes if the coin is "heads" and to answer truthfully if the coin is "tails". Suppose that respondents follow these instructions and do, in fact, answer truthfully when the coin comes up tails. Explain how you can determine the fraction of respondents who have cheated on an exam from the fraction who respond Yes to the question.

13. Suppose X and Y are independent discrete random variables. X can take on the values $0, 1, 2, 3, 4, 5$, each with probability $1/6$. Y can take on the values $10, 11, 12$, each with probability $1/3$. $Z = X + Y$.

    (a) What is the pdf of Z? (Jargon: the pdf of Z is called the *convolution* of the pdfs of X and Y.)

(b) Suppose X and Y can take on the same values as in (a). The pdf of Y is as in (a). The pdf of X is different and unknown. Suppose you know the pdf of Z. Show how to compute the pdf of X. (Jargon: This is called *deconvolution*.)

(c) Now suppose X and Y are independent and continuously distributed with densities f_X and f_Y. $Z = X + Y$. Write an expression for the pdf of Z in terms of the pdfs of X and Y.

14. Suppose that X has density $f(x)$ and $\mu = \mathbb{E}(X)$. Let m denote the median of of X, and assume $\mathbb{E}(X^2)<\infty$.

    (a) Show that $a = \mu$ solves $\min_a \mathbb{E}[(X - a)^2]$.

    (b) Show that $a = m$ solves $\min_a \mathbb{E}[|X - a|]$.

    (c) The density of X is e^x for $-\infty<x<0$ and zero elsewhere. Find the value of a that minimizes $\mathbb{E}|X-a|$.

    (d) Y and X are two random variables.

        i. Find the values of α_0 and α_1 that minimize $MSE(\alpha_0, \alpha_1) = \mathbb{E}[(Y-\alpha_0-\alpha_1 X)^2]$. (Assume the existence of any necessary moments.)

        ii. Find the values of α_0, α_1, and α_2 that minimize $MSE(\alpha_0, \alpha_1, \alpha_2) = \mathbb{E}[(Y-\alpha_0-\alpha_1 X-\alpha_2 X^2)^2]$.

15. Let Y be the number of successes in n independent Bernouli trials with $p = 0.25$. Find the smallest value of n such that $\mathbb{P}(1 \leq Y) \geq 0.7$.

16. Suppose that X and Y are independent random variables, g and h are two functions, and $\mathbb{E}(g(X))$ and $\mathbb{E}(h(Y))$ exist. Show $\mathbb{E}(g(X)h(Y)) = \mathbb{E}(g(X)) \times \mathbb{E}(h(Y))$.

17. $X \sim \mathcal{U}[1, 2]$, that is X is distributed uniformly between 1 and 2. Conditional on $X = x$, $Y \sim \mathcal{U}[x, 2 + x]$.

    (a) Derive $\mathbb{E}(X)$, $\mathbb{E}(X^2)$ and σ_X^2.

    (b) Derive $\mathbb{E}(Y)$, $\mathbb{E}(Y^2)$ and σ_Y^2.

18. X and Y are two random variables with finite mean and variance, but may not be independent.

    (a) Show that the covariance between X and Y is finite.

    (b) Suppose X and Y are independent. Show that the covariance and correlation are equal to zero.

    (c) Construct of example in which X and Y are uncorrelated, but are not independent.

19. Let X and Y denote two random variables with joint distributions $f(x, y)$. Show that $\mathrm{Var}(Y) = \mathbb{E}(\mathrm{Var}(Y|X)) + \mathrm{Var}(\mathbb{E}(Y|X))$. This is sometimes called the "law of total variation".

20. Suppose X has probability density function $f(x) = 1/x^2$ for $x \geq 1$, and $f(x) = 0$ elsewhere.

    (a) Compute $\mathbb{P}(X \geq 5)$.

    (b) Derive the CDF of X.

    (c) Show that the mean of X does not exist.

    (d) Let $Y = 1/X$. Derive the probability density function of Y.

    (e) Compute $\mathbb{E}(Y)$.

21. Let $M(t)$ denote the moment generating function of X. Let $S(t) = \ln(M(t))$, $S'(t)$ denote the first derivative of S, $S''(t)$ denote the second derivative, and so forth.

    (a) Show that $S'(0) = \mathbb{E}(X)$.

    (b) Show that $S''(0) = \text{Var}(X)$.

    (c) Show that $S'''(0)$ is the 3rd centered moment of X.

22. Y and X are two random variables with $Y|X = x \sim \mathcal{N}(0, x^2)$ and $X \sim \mathcal{U}[1, 2]$.

    (a) What is the joint density of Y and X?

    (b) Compute $\mathbb{E}(Y^4)$. (Hint: if $Q \sim \mathcal{N}(0, \sigma^2)$, then $E(Q^4) = 3\sigma^4$.)

    (c) What is the density of $W = \frac{Y}{X}$? (Hint: What is the density of W conditional on $X = x$?)

23. Suppose that $W \sim \chi^2_3$ and the distribution of X given $W = w$ is $\mathcal{U}[0, w]$ (that is uniform on 0 to w).

    (a) Find $\mathbb{E}(X)$.

    (b) Find $\text{Var}(X)$.

    (c) Suppose $W = 4$. What is the value of the minimum mean square error forecast of X?

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

25. Let $X \sim \mathcal{N}(0, 1)$ and suppose that

$$Y = \left\{ \begin{array}{c} X \text{ if } 0.5 < X < 1 \\ 0 \text{ otherwise} \end{array} \right\}$$

    (a) Find the distribution of Y.

    (b) Find $\mathbb{E}(Y)$.

    (c) Find $\text{Var}(Y)$.

26. $X \sim \mathcal{N}(0, 1)$ and $Y|X = x$ is distributed $\mathcal{N}(x, 1)$.

    (a) Find

        i. $\mathbb{E}(Y)$

        ii. $\mathbb{E}(Y^2)$

        iii. $\mathbb{E}(XY)$

        iv. $\text{Var}(Y)$

        v. $\text{Cov}(X, Y)$ (You might find the law of iterated expectations to be useful.)

    (b) Show that the joint distribution of (X, Y) is multivariate normal.

27. Complete the following calculations

    (a) Suppose $X \sim \mathcal{N}(5, 100)$. Find $\mathbb{P}(-2 < X < 7)$.

    (b) Suppose $X \sim \mathcal{N}(50, 4)$. Find $\mathbb{P}(X \geq 48)$

    (c) Suppose $X \sim \chi^2_4$. Find $\mathbb{P}(X \geq 9.49)$

(d) Suppose $X \sim \chi^2_6$. Find $\mathbb{P}(1.64 \leq X \leq 12.59)$

(e) Suppose $X \sim \chi^2_6$. Find $\mathbb{P}(X \leq 3.48)$

(f) Suppose$X \sim t_9$. Find $\mathbb{P}(1.38 \leq X \leq 2.822)$

(g) Suppose $X \sim \mathcal{N}(5, 100)$. Find $\mathbb{P}(-1 < X < 6)$.

(h) Suppose $X \sim \mathcal{N}(46, 2)$. Find $\mathbb{P}(X \geq 48)$

(i) Suppose $X \sim \chi^2_6$. Find $\mathbb{P}(1.24 \leq X \leq 12.59)$

(j) Suppose $X \sim F_{3.5}$. Find $\mathbb{P}(X > 7.76)$

(k) Suppose $X \sim t_{10}$. Find $\mathbb{P}(X \leq 2.228)$

(l) Suppose $X \sim \mathcal{N}(1, 4)$. Find $\mathbb{P}(1 < X < 9)$.

28. Find the 90th percentile of the $\mathcal{N}(10, 30)$ distribution.

29. Suppose $X \sim \mathcal{N}(\mu, \sigma^2)$ and $\mathbb{P}(X < 89) = .90$ and $\mathbb{P}(X < 94) = .95$. Find μ and σ^2.

30. Suppose $X \sim \mathcal{N}(\mu, \sigma^2)$, show that $\mathbb{E}(|X - \mu|) = \sigma\sqrt{2/\pi}$.

31. Suppose that a random variable X is distributed $\mathcal{N}(6, 2)$. Let ε be another random variable that is distributed $\mathcal{N}(0, 3)$. Suppose that X and ε are statistically independent. Let the random variable Y be related to X and ε by $Y = 4 + 6X + \varepsilon$.

    (a) Show that the joint distribution of (X, Y) is normal. Include in your answer the value of the mean vector and covariance matrix for the distribution.

    (b) Suppose that you knew that $X = 10$. Find the optimal (minimum mean square error) forecast of Y.

32. $Z \sim \mathcal{N}(0, 1)$.

    (a) Derive the first 4 moments of Z. (Hint: use the MGF.)

    (b) $W \sim \chi^2_1$. Use the result in (a) to derive the mean and variance of W.

    (c) $W \sim \chi^2_2$. Use the result in (b) to derive the mean and variance of W.

    (d) $W \sim \chi^2_n$. Use the result in (b) to derive the mean and variance of W.

33. Suppose that X and Y are two random variables with a joint normal distribution. Further suppose $\text{Var}(X) = \text{Var}(Y)$. Let $U = X + Y$ and $V = X - Y$.

    (a) Show that U and V are jointly normally distributed.

    (b) Show that U and V are independent.

34. $Y \sim \mathcal{N}(10, 4)$. Find the shortest interval of the real line, say C, such that $\mathbb{P}(Y \in C) = 0.90$.

35. Suppose Y has a continuous pdf f. How would you find values a and b such that (i) $\mathbb{P}(a \leq Y \leq b) = 0.90$ and (ii) $b - a$ is minimized?

36. $X \sim \mathcal{N}(0, 1)$ and $Y|X$ is $\mathcal{N}(X, 1)$.

    (a) Use the law of iterated expectations to find:

i. $\mathbb{E}(Y)$

ii. $\mathbb{E}(Y^2)$

iii. $\mathbb{E}(XY)$

(b) Using the results in (a), find

i. $\text{Var}(Y)$

ii. $\text{Cov}(X, Y)$

37. Show that if X is distributed $\mathcal{N}(0, 1)$ and g is an integrable function with first derivative g'.

(a) Show that $\mathbb{E}(g'(X) - Xg(X)) = 0$. (Hint: use integration by parts.)

(b) Use this result to show $E(X^{i+1}) = iE(X^{i-1})$.

38. $X \sim \mathcal{N}(0, 1)$. The distribution of $Y|X = x$ *is* $\mathcal{N}(6 + 2x, 1)$.

(a) (a) Prove that the joint distribution of X and Y is normal.

(b) (b) Derive the marginal distribution of Y.

(c) (c) Derive the distribution of $X|Y = y$.

39. Let $X \sim N_p(\mu, \Sigma)$. Partition X as $X = (X_1^{'}, X_2^{'})^{'}$, $\mu = (\mu_1^{'}, \mu_2^{'})^{'}$, and

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

such that X_1 and μ_1 are k–dimensional and Σ_{11} is a $k \times k$ matrix. Show that $\Sigma_{12} = 0$ implies that X_1 and X_2 are independent.

40. Let l denote a $p \times 1$ vector of 1's and I_p denote the identity matrix of order p. Suppose the $p \times 1$ vector Y is distributed as $Y \sim \mathcal{N}(0, \sigma^2 I_p)$, with $\sigma^2 = 4$, and let $M = I_p - (1/p)ll'$.

(a) Show that M is idempotent.

(b) Compute $\mathbb{E}(Y'MY)$.

(c) $\text{Var}(Y'MY)$.

(d) Suppose $\mathbb{E}(Y) = 0$ and $\text{Var}(Y) = \sigma^2 I_p$, but Y is not normally distributed. Would your answer to (b) change? Would your answer to (c) change? Explain.

41. Suppose that X and Y are two random vectors that satisfy:

$$E \begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} \mu_X \\ \mu_Y \end{bmatrix} \text{and Var} \begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{XY} & \Sigma_{YY} \end{bmatrix}$$

where X is $k \times 1$ and Y is scalar. Let $\hat{Y} = a + bX$.

(a) Derive the values of the scalar a and the $1 \times k$ matrix b that minimize $\mathbb{E}[(Y - \hat{Y})^2]$.

(b) Find an expression for the minimized value of $\mathbb{E}[(Y - \hat{Y})^2]$.

42. Suppose that the 3×1 vector X is distributed $\mathcal{N}(0, I_3)$, and let V be a 2×3 non-random matrix that satisfies $VV' = I_2$. Let $Y = VX$.

(a) Show that $Y'Y \sim \chi^2_2$

(b) Use Chebyshev's inequality to construct an upper bound for $\mathbb{P}(Y'Y>6)$.

(c) Find the exact value of $\mathbb{P}(Y'Y>6)$.

43. Suppose that $W \sim \chi^2_3$ and that the distribution of X given $W = w$ is $\mathcal{U}[0, w]$ (that is uniform on 0 to w).

    (a) Find $\mathbb{E}(X)$.

    (b) Find $\text{Var}(X)$.

    (c) Suppose $W = 4$. What is the value of the minimum mean square error forecast of X?

44. X and Y are random variables. The marginal distribution of X is χ^2_1. The conditional distribution of $Y|X = x$ is $\mathcal{N}(x, x)$.

    (a) Derive the mean and variance of Y.

    (b) Suppose that you are told that $X = 1$. Derive the minimum mean square error forecast of Y.

    (c) (More difficult) Suppose that you are told that $X<1$. Derive the minimum mean square error forecast of Y.

45. $X \sim \mathcal{N}(1, 4)$ and $Y = e^X$.

    (a) What is the density of Y?

    (b) Use Jensen's inequality to show that $\mathbb{E}(Y) \geq e^{\mathbb{E}(X)}$.

    (c) Compute $\mathbb{E}(Y)$. (Hint: What is the MGF for X?) Is the inequality in (b) strict?

46. Suppose X is a random variable that can take on the values $0, 1, 2, ...$ with

$$f_X(x) = \frac{m^x e^{-m}}{x!}$$

    (This is the Poisson pdf). You can verify that the mean and variance of X are both equal to m. Suppose $\mathbb{E}(X) = 80$.

    (a) Use Chebyshev's inequality to find a lower bound for $\mathbb{P}(70 \leq X \leq 90)$.

    (b) Is the answer the same for $\mathbb{P}(70 < X < 90)$? Explain.

47. Suppose X_i, $i = 1, ..., n$ are a sequence of *iid* random variables with $\mathbb{E}(X_i) = 0$ and $\text{Var}(X_i) = \sigma^2 < \infty$. Let $Y_i = iX_i$ and $\overline{Y} = n^{-2} \sum_{i=1}^n Y$. Prove $\overline{Y} \xrightarrow{p} 0$.

48. Let $\overline{X}_1$ and $\overline{X}_2$ denote the sample means from two independent randomly drawn samples of Bernoulli random variables. Both of the samples have size n and the Bernoulli population has $\mathbb{P}(X = 1) = 0.5$. Use an appropriate approximation to determine how large n must be so that $\mathbb{P}(|\overline{X}_1 - \overline{X}_2| < .01) = .95$.

49. Three candidates are running for the same elected position. Call these candidates, A, B, and C. A random sample of n voters is collected. Each voter is asked to state a preference for one (and only one) of three candidates. Let p_A denote the fraction of voters in the population who prefer A and p_B

the fraction that prefer B. (Note $p_C = 1 - p_A - p_B$.) Similarly let $\hat{p}_A$ and $\hat{p}_B$ denote the fractions in the sample. Show that

$$\sqrt{n}\begin{pmatrix} \hat{p}_A - p_A \\ \hat{p}_B - p_B \end{pmatrix} \Rightarrow \mathcal{N}(0, V)$$

and derive an expression for V.

50. Suppose $W \sim F_{p,q}$.

    (a) Suppose that $p = 20$ and $q = 40$. Find $\mathbb{P}(W > 1.6)$.

    (b) Let $q = 2p$. Prove that (as p grows large) $\sqrt{p}(W - 1) \Rightarrow \mathcal{N}(0, V)$ and derive an expression for V. (Hint: Write out the definition of an $F_{p,2p}$ random variable.)

    (c) Use your answer in (b) to compute an approximation to $\mathbb{P}(W > 1.6)$.

51. Suppose $\epsilon_i \sim iid$ with mean 1 and variance 2. Let $\bar{\varepsilon} = \frac{1}{n}\sum_{i=1}^{n} \varepsilon_i$.

    (a) Show that $\sqrt{n}(\bar{\varepsilon} - 1) \Rightarrow \mathcal{N}(0, V)$ and derive an expression for V.

    (b) Suppose $n = 50$. Use your result in (a) to derive an approximation for $\mathbb{P}(\bar{\varepsilon} > 1.2)$.

    (c) Suppose $\epsilon_i \sim \mathcal{N}(1, 2)$. How good is the approximation in (b)? (Feel free to use a computer)

    (d) Suppose $\epsilon_i \sim \chi^2_1$. How good is the approximation in (b)? (Feel free to use a computer)

52. $X_i \sim \mathcal{U}[0, 1]$ for $i = 1, \ldots, n$. Let $Y = \sum_{i=1}^{n} X_i$. Suppose that $n = 100$. Use an appropriate approximation to estimate the value of $\mathbb{P}(Y \leq 52)$.

53. $X \sim \mathcal{N}(3, 10)$, $Y \sim \mathcal{N}(3, 10)$, X and Y are independent, and $X_n = X + 1/n$.

    (a) Show that $X_n \xrightarrow{p} X$.

    (b) Show that $X_n \Rightarrow X$.

    (c) Show that $X_n \Rightarrow Y$.

    (d) Does $X_n \xrightarrow{p} Y$? Explain.

54. $X \sim \mathcal{N}(3, 10)$, $Y \sim \mathcal{N}(3, 10)$, X and Y are independent, and $X_n = (1 + n^{-1})X$.

    (a) Show that $X_n \xrightarrow{p} X$.

    (b) Show that $X_n \Rightarrow X$.

    (c) Show that $X_n \Rightarrow Y$.

    (d) Does $X_n \xrightarrow{p} Y$? Explain.

55. Suppose $y_{it} = \mu + \alpha_i \gamma_t + \epsilon_{it}$ for $i = 1, ..., n$ and $t = 1, ..., T$. Write

$$\bar{y} = \frac{1}{nT}\sum_{i,t} y_{it} = \mu + \left(\frac{1}{n}\sum_i \alpha_i\right)\left(\frac{1}{T}\sum_t \gamma_t\right) + \frac{1}{nT}\sum_{i,t} \epsilon_{it}$$

Suppose $\alpha_i, \gamma_t, \epsilon_{it}$ are iid with means zero. Show

$$\sqrt{nT}(\bar{y} - \mu) \Rightarrow \sigma_\alpha Z_\alpha \sigma_\gamma Z_\gamma + \sigma_\epsilon Z_\epsilon$$

where the Zs are *iid* $\mathcal{N}(0, 1)$ random variables.

56. Suppose $X_t = \varepsilon_t + \varepsilon_{t-1}$, where ε_t are *i.i.d.* with mean 0 and variance 1.

   (a) Find an upper bound for $\mathbb{P}(|X_t| \geq 3)$.

   (b) Show that $\frac{1}{n}\sum_{t=1}^{n} X_t \xrightarrow{p} 0$.

   (c) Show that $\frac{1}{\sqrt{n}}\sum_{t=1}^{n} X_t \Rightarrow \mathcal{N}(0, V)$, and derive an expression for V.

   (d) Now suppose that X_t follows the "MA(q)" process $X_t = \varepsilon_t - \sum_{i=1}^{q} \theta_i \varepsilon_{t-i}$, where θ_i are finite constants. Repeat parts (b) and (c).

57. Let $Y \sim \chi^2_n$, let $X_n = n^{-1/2}Y - n^{1/2}$, and $W_n = n^{-1}Y$.

   (a) Show that $X_n \Rightarrow X \sim \mathcal{N}(0, 2)$.

   (b) Show that $W_n \xrightarrow{p} 1$.

   (c) Suppose that $n = 30$. Calculate $\mathbb{P}(Y > 43.8)$.

   (d) Use the result in (a) to approximate $\mathbb{P}(Y > 43.8)$.

58. Suppose that X_i are $\mathcal{U}[0, 1]$ for $i = 1, ..., n$. Let $s^2 = (n-1)^{-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$, where $\bar{X}_n = n^{-1}\sum_{i=1}^{n} X_i$, and $s = \sqrt{s^2}$.

   (a) Show $\mathbb{E}(s^2) = 1/12$.

   (b) Show $s^2 \xrightarrow{p} 1/12$.

   (c) Show $s \xrightarrow{p} \sqrt{1/12}$

   (d) Is $\mathbb{E}(s) = \sqrt{1/12}$? $\mathbb{E}(s) > \sqrt{1/12}$? $\mathbb{E}(s) < \sqrt{1/12}$? Explain.

59. $Z \sim \mathcal{N}(0, 1)$.

   (a) Derive the first 4 moments of Z.

   (b) $W \sim \chi^2_{10}$. Use the result in (a) to derive the mean and variance of W.

   (c) Let $W_i \sim i.i.d. \chi^2_{10}$. Let $\overline{W} = \frac{1}{n}\sum_{i=1}^{n} W_i$

      i. Show that $\mathbb{E}(\overline{W}) = 10$.

      ii. Show that $\overline{W} \xrightarrow{p} 10$.

      iii. Show that $\frac{1}{n}\sum_{i=1}^{n}(W_i - \overline{W})^2 \xrightarrow{p} \sigma^2_W$.

      iv. Does $\mathbf{E}\left[\frac{1}{n}\sum_{i=1}^{n}(W_i - \overline{W})^2\right] = \sigma^2_W$?

60. $Z \sim \mathcal{N}(0, 1)$

   (a) Prove that $\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(Z_i^2 - 1) \Rightarrow \mathcal{N}(0, V)$ and determine the value of V.

   (b) Use the result in (a) to find an approximation to $\mathbb{P}\left(\sum_{i=1}^{30} Z_i^2 < 40\right)$.

   (c) Find the exact value for $\mathbb{P}\left(\sum_{i=1}^{30} Z_i^2 < 40\right)$.

61. Suppose that $X_i \sim i.i.d. \mathcal{N}(0, \Sigma)$ for $i = 1, ..., n$, where X_i is a $p \times 1$ vector. Let α denote a non-zero and non-stochastic $p \times 1$ vector and let

$$Y_n = \frac{\alpha' X_n X_n' \alpha}{\frac{1}{n-1}\sum_{i=1}^{n-1} \alpha' X_i X_i' \alpha}$$

(a) Prove that $Y_n \sim F_{1,n-1}$.

(b) Prove that $Y_n \Rightarrow Y \sim \chi^2_1 \quad (\text{as } n \to \infty)$.

62. Let $Y \sim N_p(0, \lambda I_p)$, where λ is a positive scalar. Let Q denote a $p \times k$ non-stochastic matrix with full column rank, $P_Q = Q(Q^{'}Q)^{-1}Q^{'}$ and $M_Q = I - P_Q$. Finally, Let $X_1 = Y^{'}P_QY$ and $X_2 = Y^{'}M_QY$. (A few useful matrix facts: (i) If A and B are two matrices with dimensions so that AB and BA are defined, then $tr(AB) = tr(BA)$; (ii) if Q is an idempotent matrix then $rank(Q) = tr(Q)$.)

(a) Prove that $\mathbb{E}(X_2) = (p-k)\lambda$.

(b) Let $W = (X_1/X_2)((p-k)/k)$. Show that $W \sim F_{k,p-k}$.

(c) Suppose that $p = 30$ and $k = 15$. Find $\mathbb{P}(W > 2.4)$.

(d) Suppose that $k = [p/2]$ where $[z]$ denote the greatest integeter less than or equal to z. Show that as $p \to \infty$, $\sqrt{p}(W-1) \Rightarrow R \sim \mathcal{N}(0, V)$, and derive an expression for V.

(e) Use your answer in (d) to approximate the answer to (c).

63. $X_i \sim iid$ with mean 0 and variance σ^2. Consider the method of moments estimator $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n X_i^2$.

(a) Show that $\sqrt{n}(\hat{\sigma}^2 - \sigma^2) \Rightarrow \mathcal{N}(0, V)$ and derive an expression for V. (You will have to make additional assumptions about the existence of moments. Clearly state your assumptions.)

(b) Let $\sqrt{\hat{\sigma}^2} = \hat{\sigma}$ denote an estimator of the standard deviation. Show that $\sqrt{n}(\hat{\sigma} - \sigma) \Rightarrow \mathcal{N}(0, W)$ and derive an expression for W.

(c) Propose estimators for V and W and prove that your proposed estimators are consistent.

64. X_i are distributed iid Bernouli, with $\mathbb{E}(X_i) = p$, for $i = 1, \ldots, n$. Let$\bar{X}$ denote the sample mean.

(a) Prove that $\sqrt{n}(\bar{X} - p) \Rightarrow \mathcal{N}(0, V)$ and derive an expression for V.

(b) Prove that $\sqrt{n}(\mathcal{N}(\bar{X}) - \mathcal{N}(p)) \Rightarrow \mathcal{N}(0, W)$ and derive an expression for W.

65. Assume that $Y_i \sim iid\mathcal{N}(\mu, 1)$ for $i = 1, \ldots, n$. Let $\hat{\mu} = \mathbf{1}(|\bar{Y}| > c)\bar{Y}$, where c is a positive constant. (The notation $\mathbf{1}(|\bar{Y}| > c)$ means that $\mathbf{1}(|\bar{Y}| > c) = 1$ if $|\bar{Y}| > c$, and is equal to 0 otherwise).

(a) Derive $\mathbb{E}(\hat{\mu})$.

(b) Discuss the consistency of $\hat{\mu}$ when

i. $|\mu| > c$

ii. $|\mu| < c$

iii. $|\mu| = c$.

66. $X_i \sim iid\mathcal{U}[0, \theta]$, for $i = 1, \ldots, n$. Let $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$.

(a) Show that $2\bar{X} \xrightarrow{p} \theta$.

(b) Show that $\sqrt{n}(2\bar{X} - \theta) \Rightarrow N0, V)$ and derive an expression for V.

(c) Derive the MLE of θ.

(d) Derive the CDF of $\hat{\theta}^{MLE}$.

(e) Prove that $\hat{\theta}^{MLE} \xrightarrow{p} \theta$.

67. Suppose that $X_i \sim iid\mathcal{N}(0, \sigma^2)$, $i = 1, ..., n$.

   (a) Construct the MLE of σ^2

   (b) Show that the estimator is unbiased.

   (c) Compute the variance of the estimator.

   (d) Show that $\sqrt{n}(\hat{\sigma}^2_{MLE} - \sigma^2) \Rightarrow \mathcal{N}(0, V)$. Derive and expression for V.

   (e) Construct the MLE of σ.

   (f) Is the MLE of σ unbiased? Explain.

   (g) Show that $\sqrt{n}(\hat{\sigma}_{MLE} - \sigma) \Rightarrow \mathcal{N}(0, W)$. Derive and expression for W.

   (h) Suppose that $X_i \sim iid(0, \sigma^2)$ with $\mathbb{E}(X_i^4) = \kappa < \infty$. (Note, Normality is not assumed.) Show that the MLEs computed in (a) and (e) (that is, the estimators that assumed normality) are consistent for σ^2 and σ, respectively. Repeat parts (d) and (g). What changes?

68. Suppose the $X_i \sim iid\mathcal{N}(\mu, \sigma^2)$, $i = 1, \ldots, n$. Let the 2×1 vector θ be given by $\theta = (\mu \ \sigma^2)'$.

   (a) Construct the MLE of θ.

   (b) Is the estimator unbiased?

   (c) Show that $\sqrt{n}(\hat{\theta}_{MLE} - \theta) \Rightarrow \mathcal{N}(0, V)$. Derive an expression for V.

69. $Y_i \sim iid\mathcal{N}(\mu, 1)$ for $i = 1, 3, 5, \ldots n$–1, $Y_i \sim i.i.d.\mathcal{N}(\mu, 4)$ for $i = 2, 4, \ldots, n$ where n is even, and Y_i and Y_j are independent for i odd and j even.

   (a) Write the log-likelihood function.

   (b) Derive the MLE of μ.

   (c) Is the MLE preferred to $\bar{Y}$ as an estimator of μ ? Explain. (Discuss the properties of the estimators. Are the estimators unbiased? Consistent? Asymptotically Normal? Which estimator has lower quadratic risk?)

70. (Y, X) are two random variables. The distribution of X is Bernoulli with parameter p. The distribution of $Y|X = x$ is $\mathcal{N}(\mu(x), \sigma(x)^2)$

   (a) Suppose $\sigma(0)^2 = \sigma(1)^2 = 1$, $\mu(0) = 0$, $\mu(1) = 3$, and $p = 0.6$.

      i. Compute $\mathbb{P}(Y \leq 1)$.

      ii. You observe $Y = 1$.

         A. What is the distribution of $X|Y = 1$.

         B. You need to guess the value of X for a decision. Let g denote your guess, suppose your loss is $(g - X)^2$. What value of g should you guess? Why?

   (b) Now suppose $\sigma(0)^2 = 2$, $\sigma(1)^2 = 4$, $\mu(0) = -1$, $\mu(1) = 3$, and $p = 0.6$. Compute the variance of Y.

   (c) (Y_i, X_i), for $i = 1, \ldots, n$ are *i.i.d.* draws from the distribution, and $\sigma(0)$, $\sigma(1)$, $\mu(0)$, $\mu(1)$, and p and unknown.

      i. Write the log-likelihood for $\sigma(0)$, $\sigma(1)$, $\mu(0)$, $\mu(1)$, and p.

ii. Show that the MLE of $\mu(1)$ can be written as

$$\widehat{\mu(1)} = \left(\frac{n}{\sum_{i=1}^{n} X_i}\right) \frac{1}{n} \sum_{i=1}^{n} X_i Y_i$$

iii. Show that $\widehat{\mu(1)} \xrightarrow{p} \mu(1)$.

iv. Show that $\sqrt{n}\left(\widehat{\mu(1)} - \mu(1)\right) \Rightarrow \mathcal{N}(0, W)$ and derive an expression for W.

v. Show that $\sum_{i=1}^{n} X_i$ is sufficient for p.

71. $X_i \sim iid$ Bernoulli with parameter p, for $i = 1, \ldots, n$. Let $X_{1:n} = (X_1, X_2, ...X_n)'$.

    (a) Derive the likelihood function $f(X_{1:n}|p)$.

    (b) Derive the score, $S(X_{1:n}, p)$ and show that it has mean zero.

    (c) Let $\hat{p} = \bar{X}$. Show that $\hat{p}$ is the "best unbiased estimator" of p. ("Best" means minimum mean square estimator in this context.)

72. Suppose X is a Bernoulli random variable with $\mathbb{P}(X = 1) = p$. Suppose that you have a sample of i.i.d. random variables from this distribution, $X_1, X_2, ..., X_n$. Let p_o denote the true, but unknown value of p.

    (a) Write the log-likelihood function for p.

    (b) Compute the score function for p and show that the score has expected value 0 when evaluated at p_o.

    (c) Compute the information $\mathcal{I}$.

    (d) Show the maximum likelihood estimator is given by $\hat{p} = \overline{X}$.

    (e) Show that $\hat{p}$ is unbiased.

    (f) Show that $\sqrt{n}(\hat{p} - p_o) \xrightarrow{d} \mathcal{N}(0, V)$, and derive an expression for V.

    (g) Propose a consistent estimator for V.

73. Let $X_i, i = 1, ..., n$ denote an i.i.d. sample of Bernoulli random variables with $\mathbb{P}(X_i = 1) = p$. I am interested in a parameter θ that is given by $\theta = e^p$.

    (a) Construct the likelihood function for θ.

    (b) Show that $\hat{\theta}_{MLE} = e^{\overline{X}}$.

    (c) Show that $\hat{\theta}_{MLE}$ is consistent.

    (d) Show that $\sqrt{n}(\hat{\theta}_{MLE} - \theta) \Rightarrow \mathcal{N}(0, V)$, and derive an expression for V.

74. Suppose $\sqrt{n}(\hat{\theta} - \theta_0) \Rightarrow \mathcal{N}_k(0, V)$. Let $\hat{V}$ be a consistent estimator of V.

    (a) Show that $(\hat{\theta} - \theta_0)' \left(\frac{1}{n}\hat{V}\right)^{-1} (\hat{\theta} - \theta_0) \Rightarrow \chi^2_k$.

75. Suppose that a test statistic, T_n, for testing $H_o : \theta = 0$ vs. $H_a : \theta \neq 0$ constructed from a sample of size n is known to have a certain asymptotic null distribution

$$T_n \Rightarrow X \sim F$$

where X is a random variable with distribution function $F(x)$, and convergence obtains under the null hypothesis. Suppose that the distribution $F(x)$ depends on an unknown parameter ρ. Represent this dependence by $F(x, \rho)$. Let $x(\alpha, \rho)$ solve $F(x, \rho) = \alpha$, and suppose that $x(\alpha, \rho)$ is continuous in ρ. Let ρ_o denote the true value of ρ and let $\hat{\rho}_n$ denote a consistent estimator of ρ. Prove that a critical region defined by

$$T_n > x(\alpha, \hat{\rho}_n)$$

produces a test with asymptotic size equal to α. (Such a test is said to be asymptotically similar.)

76. Suppose the $Y_i \sim iid\, \mathcal{N}(\mu, \sigma^2)$, $i = 1, \ldots, n$. Let the 2×1 vector θ be given by $\theta = (\mu\ \sigma)'$.

    (a) Construct the MLE of θ.

    (b) Show that $\sqrt{n}(\hat{\theta}_{MLE} - \theta) \Rightarrow \mathcal{N}(0, V)$. Derive an expression for V.

    (c) Propose a consistent estimator for V. Prove that your estimator is consistent.

    (d) In a sample with $n = 100$, $\sum_{i=1}^{n} Y_i = 1049$, and $\sum_{i=1}^{n} Y_i^2 = 13149$.

        i. What is the value of $\hat{\theta}_{MLE}$?

        ii. Use the Wald statistic (constructed using your consistent estimator for V) to test the null hypothesis that $\mu = 10$ and $\sigma = 4$. Set the size of the test equal to 5%.

77. $Y_i \sim iid\, N(0, \sigma^2)$, for $i = 1, \ldots, n$. Let $s^2 = \frac{1}{n}\sum_{i=1}^{n} Y_i^2$. I am interested in the competing hypotheses $H_o : \sigma^2 = 1$ versus $H_a : \sigma^2 > 1$. Suppose $n = 15$, and I decide to reject the null if $s^2 > 5/3$.

    (a) What is the size of the test?

    (b) Suppose that the true value of σ^2 is 1.5. What is the power of the test?

    (c) Is this the most powerful test? Explain.

78. Y_i for $i = 1, \ldots, n$ are independent and distributed $\mathcal{N}(0, \sigma_i^2)$, where $\sigma_i^2 = 1 + g/i$, and g is non-negative constant. I am interested in the competing hypotheses $H_o : g = 0$ versus $H_a : g = 1$.

    (a) Derive the optimal test and provide explicit instructions for carrying out the optimal test with size = 10%.

    (b) Is the test you derived in (a) the uniformly most powerful test for the composite alternative H_a: $g > 0$?

79. $Y_i \sim iid$ Bernoulli with parameter p, for $i = 1, \ldots, n$. Show that $S = \sum_{i=1}^{n} Y_i$ is a sufficient statistic for p.

80. A researcher is interested in computing the likelihood ratio statistic

$$LR(Y) = \frac{f(Y, \theta_1)}{f(Y, \theta_2)}$$

but finds the calculation difficult because the density function $f(.)$ is very complicated. However, it easy to compute the joint density $g(y, x, \theta)$ where X denotes another random variable. The researcher has data on Y, but not on X. Show that the researcher can compute $LR(Y)$ using the formula

$$LR(Y) = \frac{f(Y, \theta_1)}{f(Y, \theta_2)} = \mathbb{E}_{\theta_2}\left[\frac{g(Y, X, \theta_1)}{g(Y, X, \theta_2)} | Y\right]$$

where the notation $\mathbb{E}_{\theta_2}[\bullet\,|Y]$ means that the expectation should be taken using the distribution conditional on Y and using the parameter value θ_2.

81. Suppose the $Y_i \sim iid\, N(0, \sigma^2)$, for $i = 1, \ldots, n$.

    (a) Suppose the MLE of σ^2 is $\hat{\sigma}^2_{MLE} = 4$ and $n = 100$. Test $H_o : \sigma^2 = 3$ versus $H_a : \sigma^2 \neq 3$ using a Wald test with a size of 10%

    (b) Construct a 95% confidence interval for σ^2.

    (c) Suppose the MLE of σ^2 is $\hat{\sigma}^2_{MLE} = 4$ and $n = 100$. Test $H_o : \sigma^2 = 3$ versus $H_a : \sigma^2 = 4$ using a test of your choosing. Set the size of the test at 5%. Justify your choice of test.

    (d) Suppose that the random variables were *iid* but non-normal. Would the confidence interval that you constructed in (b) have the correct coverage probability (*i.e.,* would it contain the true value of σ^2 with probability 0.95)? How could you modify the procedure to make it robust to the assumption of normality (i.e., so that it provided the correct coverage probability even if the data were drawn from a non-normal distribution.)

82. Suppose that Y_i is *iid* with mean μ, variance σ^2. A random sample of 100 observations are collected, and you are given the following summary statistics: $\bar{Y} = 1.23$ and $s^2 = 4.33$, where $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2$ is the sample variance.

    (a) Prove that s^2 is consistent for σ^2.

    (b) Construct an approximate 95% confidence interval for μ. Why is this an "approximate" confidence interval?

    (c) Let $\alpha = \mu^2$. Construct an approximate 95% confidence interval for α.

83. An economist is flying to a conference where he will present a new research paper. While looking over his presentation he realizes that he failed to calculate the "p-value" for the key t-statistic in his paper. Unfortunately he doesn't have any statistical tables with him, but he does have a laptop computer and a program that can generate *iid* $\mathcal{N}(0, 1)$ random variables. He decides to estimate the p-value by simulating t-distributed random variables by transforming the standard normals generated by his random number generator. Specifically, he does the following. Let $\bar{t}$ denote the numerical value of the economist's t-statistic. The p-value that he wants to estimate is

$$P = \mathbb{P}(t_{df} > \bar{t})$$

where t_{df} is a t-distributed random variable with degrees of freedom equal to df. For his problem, the correct value of df is $df = 5$. The economist generates 100 draws from the t_5 distributions and estimates P by the sample fraction of the draws that were larger than $\bar{t}$. That is, letting x_i denote the ith draw from the t_5 distribution, and q_i denote the indicator variable

$$q_i = \left\{ \begin{array}{c} 1 \text{ if } x_i > \bar{t} \\ 0 \text{ otherwise} \end{array} \right\}$$

then P is estimated as

$$\widehat{P} = \frac{1}{100}\sum_{i=1}^{100} q_i$$

(a) If you had a random number generator that only generated $i.i.d.\mathcal{N}(0,1)$ random variables, how would you generate a draw from the t_5 distribution?

(b) Prove that q_i are distributed as $i.i.d.$ Bernoulli random variables with parameter P.

(c) Compute the mean and variance of $\hat{P}$ as a function of P.

(d) Prove that $\hat{P}$ is consistent estimator or P. What sample size needs to get large for consistency?

(e) Suppose $\hat{P} = 0.04$. Construct an approximate 90% confidence interval for P. (Hint: Since the number of simulations is large ($n = 100$), you can use the Central Limit Theorem.)

(f) Write the likelihood function for P as a function of the sample values of q_i.

(g) Show that $\hat{P}$ is the MLE of P.

(h) The economist thinks that he remembers looking up the p-value for the statistic and seems to remember that the correct value was 0.03. How would you construct the most powerful test for the null hypothesis that $P = 0.03$ versus the alternate that $P = 0.04$?

84. Suppose $Y_i \sim iid\, N(\mu, \sigma^2)$, for $i = 1, \ldots, n$. Suppose $n = 200$, $\sum_{i=1}^{n} Y_i = 1038$, and $\sum_{i=1}^{n} Y_i^2 = 7402$.

(a) Using the appropriate Wald statistic, construct a 95% confidence set for μ.

(b) Using the appropriate Wald statistic, construct a 95% confidence set for σ^2.

(c) Using the appropriate Wald statistic, construct a 95% confidence set for $(\mu,\ \sigma^2)$.

85. Suppose

$$\begin{pmatrix} X_i \\ Y_i \end{pmatrix} \sim iid\, N\left(\begin{bmatrix} \mu_X \\ \mu_Y \end{bmatrix}, I_2 \right)$$

for $i = 1, \ldots, n$. I am interested in the scalar parameter $\theta = \mu_X/\mu_Y$. Let $\hat{\theta} = \bar{X}/\bar{Y}$denote the MLE of θ.

(a) Suppose that $\mu_Y \neq 0$.

i. Show that $\sqrt{n}(\hat{\theta} - \theta) \Rightarrow W \sim \mathcal{N}(0, \Omega)$.

ii. Derive an expression for the value of Ω.

iii. Propose a consistent estimator for Ω.

(b) Suppose that $\mu_Y = 0$ and $\mu_X \neq 0$. Show that $\frac{1}{\sqrt{n}}\hat{\theta} \Rightarrow \frac{\mu_X}{Z}$ where $Z \sim \mathcal{N}(0,1)$.

(c) Suppose that $\mu_Y = \mu_X = 0$. Show that $\hat{\theta} \Rightarrow \frac{Z_1}{Z_2}$ where Z_1 and Z_2 are mutually independent $\mathcal{N}(0,1)$ random variables.

(d) Suppose that μ_Y is "small," say $\mu_Y = m_Y/\sqrt{n}$ where m_Y is number like $m_Y = 1$ or 2. Use asymptotics to argue that $\frac{1}{\sqrt{n}}\hat{\theta}$ has approximately the same distribution as $\mu_X/(m_Y + Z)$, where Z is $\mathcal{N}(0,1)$.

(e) Suppose that both μ_Y and μ_X are "small," say $\mu_Y = m_Y/\sqrt{n}$ and $\mu_X = m_X/\sqrt{n}$. Use asymptotics to argue that $\hat{\theta}$ has approximately the same distribution as $(m_X + Z_1)/(m_Y + Z_2)$.

(f) Using your insights from (a) – (e) discuss the accuracy of the approximation $\hat{\theta} \overset{a}{\sim} \mathcal{N}(\theta, V)$, where $V = \Omega/n$ and how the accuracy depends on the values of μ_X and μ_Y.

(g) Consider the 95 confidence set $\hat{\theta} \pm 1.96\sqrt{\hat{\Omega}/n}$, where $\hat{\Omega}$ is the estimator you proposed in (a.iii). Discuss the validity of this confidence sets and how the validity depends on the values of μ_X and μ_Y.

(h) Suppose you are interested in testing $H_o : \theta = \theta_0$ versus $H_a : \theta \neq \theta_0$. One procedure is to use the Wald statistic based on the δ-method approximation you worked out in part (a). Another procedures uses "Fieller's method." Here's the idea: if $\theta = \theta_0$, then $\mu_X - \theta_0\mu_Y = 0$, thus the null and alternative can be recast as $H_o : \mu_X - \theta_0\mu_Y = 0$ versus $H_a : \mu_X - \theta_0\mu_Y \neq 0$. This can be tested using a Wald statistic constructed as $\xi_{Fieller}(\theta_0) = (\bar{X} - \theta_0\bar{Y})^2 / \operatorname{Var}(\bar{X} - \theta_0\bar{Y})$.

i. Compute $\operatorname{Var}(\bar{X} - \theta_0\bar{Y})$ and show that $\xi_{Fieller}(\theta_0) \sim \chi^2_1$ under H_o. Does this result depend on the values of μ_Y and μ_X (as long as the satisfy $\mu_X - \theta_0\mu_Y = 0$)?

ii. Consider the set $A_{Fieller}(\theta) = \{\theta | \xi_{Fieller}(\theta) \leq 3.84\}$. Show that $A_{Fieller}(\theta)$ is a valid 95% confidence set.

iii. Compare the confidence set $A_{Fieller}(\theta)$ to the δ-method confidence set $\hat{\theta} \pm 1.96\sqrt{\hat{\Omega}/n}$. Are they (in some relevant sense) the same when $|\mu_X|$ and $|\mu_Y|$ are large? What if $|\mu_X|$ and $|\mu_Y|$ are small?

(i) In a sample of size $n = 200$, $\bar{X} = 4.0$ and $\bar{Y} = 5.2$. Construct the δ-method and Fieller confidence sets for θ. Are they similar?

(j) In a sample of size $n = 200$, $\bar{X} = 0.040$ and $\bar{Y} = 0.052$. Construct the δ-method and Fieller confidence sets for θ. Are they similar?

86. $Y_i \sim iid\,\mathcal{N}(\mu, 1)$ for $i = 1, \ldots, n$. I am interested in the competing hypotheses $H_o : \mu = 1$ <u>or</u> $\mu = 2$ versus $H_a : \mu = 0$. (The null is composite and includes $\mu = 1$ <u>and</u> $\mu = 2$.)

(a) Consider a test that controls "Weighted Average Size" (WAS) with a weight of 1/3 on $\mu = 1$ and a weight of 2/3 on $\mu = 2$. (That is, the test's critical region, say W, satisfies $(1/3)\mathbb{P}(Y_{1:n} \in W | \mu = 1) + (2/3)\mathbb{P}(Y_{1:n} \in W | \mu = 2) = \alpha$, where α is the pre-specified WAS of the test.)

i. Derive the most-powerful test that controls WAS. Prove that your proposed test is most powerful.

(b) A <u>valid test</u> of the null must control size under both values of μ, that is must control size when $\mu = 1$ <u>and</u> when $\mu = 2$. Is the WAS test you derived in (a) a valid test?

(c) Show that the power for any valid test is bounded above by the power of the optimal WAS test that you proposed in (a).

(d) (Might be more challenging) Show that the optimal valid test is an optimal WAS test that puts weight $p = 1$ on $\mu = 1$ and weight $p = 0$ on $\mu = 2$.

87. Assume that $Y_i | \theta \sim iid\,\mathcal{N}(\theta, 1)$ for $i = 1, \ldots, n$.

(a) Suppose that $\theta \sim \mathcal{N}(\tau, \omega^2)$. Show that the posterior distribution of θ depends on $\{Y_i\}_{i=1}^n$ only through $\bar{Y}$.

(b) Suppose that $\theta \sim w$ for arbitrary prior density w. Show that the posterior distribution of θ depends on $\{Y_i\}_{i=1}^n$ only through $\bar{Y}$.

88. $X_i \sim iid$ Bernoulli with parameter p. The prior for p is

$$p = \begin{cases} 1/3 \text{ with probability } 1/5 \\ 1/2 \text{ with probability } 4/5 \end{cases}$$

A sample of size 2 yields $X_1 = 1$ and $X_2 = 0$.

(a) Derive the posterior distribution of p.

(b) Suppose loss is quadratic. Derive the Bayes estimate of p.

89. $Y_i \sim iid\, N(\mu, 1)$, for $i = 1, \ldots, 100$. A sample is selected and $\bar{Y} = 6.4$.

(a) Construct a 95% confidence set for μ. Explain in words what the set represents.

(b) Suppose $\mu \sim \mathcal{N}(0, 10000)$ is a prior for μ. Construct a 95% credible set for μ. Explain in words what the set represents.

90. X_1 and X_2 are *iid* Bernouli random variables with parameter p. The value of p is unknown, but you know that p takes one of two values, $p = 0.3$ or $p = 0.4$. Your prior is $\mathbb{P}(p = 0.3) = 0.7$ and $\mathbb{P}(p = 0.4) = 0.3$. You observe $X_1 = 1$ and $X_2 = 0$.

(a) What is the likelihood $f(X_1 = 1, X_2 = 0|p)$ as a function of p?

(b) What is the marginal likelihood$f(X_1 = 1, X_2 = 0)$?

(c) What is the posterior $f(p = 0.3|X_1 = 1, X_2 = 0)$?

(d) Loss is quadratic. What is your best guess of the value of p?

91. Assume that $Y_i|\theta \sim iid\mathcal{N}(\theta, 1)$ for $i = 1, \ldots, 4$. The prior for θ is composed of two values

$$\theta = \begin{cases} 1 \text{ with probability } 0.4 \\ 2 \text{ with probability } 0.6 \end{cases}$$

You learn $Y_1 = 1.6$, $Y_2 = 0.9$, $Y_3 = 1.8$, and $Y_4 = 2.3$.

(a) Derive the posterior for θ?

(b) Loss is quadratic: $L(\hat{\theta}, \theta) = (\hat{\theta} - \theta)^2$. What is the Bayes estimate of θ?

92. Suppose that two random variables θ and Y have joint density $f_{\theta,Y}(\tilde{\theta}, y)$ and marginal densities $w(\theta)$ and $f_Y(y)$, and suppose that the distribution of $Y|\theta = \tilde{\theta}$ is $\mathcal{N}(\tilde{\theta}, \sigma^2)$. Denote this conditional density as $f_{Y|\theta}(y|\tilde{\theta})$.

(a) Show that $\mathbb{E}(\theta|Y = y) = \frac{\int \tilde{\theta} f_{Y|\theta}(y|\tilde{\theta}) w(\tilde{\theta}) d\tilde{\theta}}{f_Y(y)}$.

(b) Show that $\frac{\partial f_{Y|\theta}(y|\tilde{\theta})}{\partial y} = \frac{1}{\sigma^2}(\tilde{\theta} - y) f_{Y|\theta}(y|\tilde{\theta})$ and solve this for $\tilde{\theta} f_{Y|\theta}(y|\tilde{\theta})$.

(c) Show that$\frac{\partial f_Y(y)}{\partial y} = \int \frac{\partial f_{Y|\theta}(y|\tilde{\theta}) w(\tilde{\theta})}{\partial y} d\tilde{\theta}$.

(d) Using your results from (a) and (c), show that

$$\mathbb{E}(\theta|Y = y) = y + \sigma^2 S(y)$$

where $S(y) = \partial \ln(f_Y(y))/\partial y = \frac{\partial f_Y(y)/\partial y}{f_Y(y)}$. This formula provides a method for computing $\mathbb{E}(\theta|Y = y)$ when $f_{Y|\theta}(y|\tilde{\theta})$ is $N(\tilde{\theta},\ \sigma^2)$. (In the literature, this is called a "simple Bayes formula" or "Tweedie's formula" for the Bayes estimator of θ.)

93. $Y_i \sim i.i.d. \mathcal{N}(\mu, 1)$, for $i = 1, \ldots, 100$. A sample is selected and $\bar{Y} = 6.4$.

    (a) Construct a 95 confidence set for μ. Explain in words what the set represents.

    (b) Suppose $\mu \sim \mathcal{N}(0, 10000)$ is a prior for μ. Construct a 95% credible set for μ. Explain in words what the set represents.

94. Assume that $Y_i|\theta \sim iid\, \mathcal{N}(\theta, 1)$ for $i = 1, \ldots, n$. The prior for θ is $\theta \sim \mathcal{N}(\tau, \omega^2)$. Loss is quadratic.

    (a) Derive the Bayes estimator of θ?

    (b) Derive posterior risk of the Bayes estimator.

    (c) Derive the Bayes risk of the Bayes estimator.

    (d) Consider the estimator $\hat{\theta} = \bar{Y}$. Derive the Bayes risk of $\hat{\theta}$.

95. (X, Y) are two random variables with a joint distribution that depends on a scalar parameter θ; write the joint density as$f(x, y|\theta)$. Suppose that you have a prior, say $w(\theta)$, for θ.

    (a) You observe $X = x$ and $Y = y$. Write an expression for the posterior density of for θ.

    (b) Alternatively, suppose that the data come in sequentially. Specifically, you begin with the prior $w(\theta)$ and then observe $X = x$. Based on this observation you compute a posterior for θ, this posterior serves as an "updated prior", you then observe $Y = y$, and based on this observation and your "updated prior" you compute the posterior for θ. Show that this sequential procedure yields the same posterior as in (a).

96. Assume that $Y_i|\theta \sim iid\, \mathcal{N}(\theta, 1)$ for $i = 1$ and 2. Suppose loss is quadratic. Show that $\hat{\theta} = 0.25Y_1 + 0.75Y_2$ is not an admissible estimator of θ.

97. Y_1 and Y_2 are $iid\, \mathcal{N}(\mu, 1)$. Let $\hat{\mu}_1 = (Y_1 + Y_2)/2$, $\hat{\mu}_2 = (Y_1 + Y_2)/4$, and $\hat{\mu}_3 = (Y_1 + Y_2)$. Loss is quadratic: $L(\hat{\mu}, \mu) = (\hat{\mu} - \mu)^2$.

    (a) Derive the frequentist risk functions $R(\hat{\mu}_1, \mu)$, $R(\hat{\mu}_2, \mu)$, and $R(\hat{\mu}_3, \mu)$. Plot these as a function of μ.

    (b) Is $\hat{\mu}_3$ admissible? Explain.

    (c) Suppose the prior for μ is $\mathbb{P}(\mu = 0) = 0.3$ and $\mathbb{P}(\mu = 1) = 0.7$. Derive the Bayes risk $r(\hat{\mu}_1)$, $r(\hat{\mu}_2)$, and $r(\hat{\mu}_3)$. Rank the estimators.

    (d) Suppose the prior for μ is $\mu \sim \mathcal{N}(0, 1)$. Derive the Bayes risk $r(\hat{\mu}_1)$, $r(\hat{\mu}_2)$, and $r(\hat{\mu}_3)$. Rank the estimators.

    (e) Suppose the prior for μ is $\mu \sim \mathcal{N}(0, \sigma^2)$. Derive the Bayes risk $r(\hat{\mu}_1)$, $r(\hat{\mu}_2)$, and $r(\hat{\mu}_3)$. Rank the estimators.

98. Y_i for $i = 1, \ldots, n$ are independent and distributed $\mathcal{N}(0, \sigma_i^2)$, where $\sigma_i^2 = 1 + g/i$, and g is non-negative constant. (Your answers to the questions below may involve integrals. You do NOT need to solve the integrals.)

(a) I am interested in the competing hypotheses $H_o : g = 0$ versus $H_a : g = 1$. Derive the optimal test and provide explicit instructions for carrying out the optimal test with size = 10%.

(b) I have a prior, $g \sim \mathcal{U}[0, 1]$. Provide explicit instructions for computing the posterior density of g.

(c) I am interested in the competing hypotheses $H_0 : g = 0$ versus $H_a : g > 0$. Derive the test that has best average power for $0 < g \leq 1$. Provide explicit instructions for carrying out the optimal test with size = 10%.