Unit 2 Question Bank

Instructions

For the initial assignment of 10 problems please use the following parameters to select your 10:

Group	Number of Problems
1	1
2	2
3	2
4	2 or 3
5	2 or 3

Group 1

1. A marine biologist observes whether each of two tagged kelp crabs molts during a week (yes/no). Let X = number that molt. List the support of X.

2. Adult newt lengths are approximately normal with $\mu = 12.0$ cm and $\sigma = 1.5$ cm. What is the z-score of a 13.2 cm newt?

3. Each seed germinates independently with probability $p = 0.7$. For a single seed, name the distribution of success/failure and give $\mathbb{E}X$.

4. Mussel shell lengths have sample standard deviation $\sigma = 4$ mm and sample mean $\mu = 10$ mm. Given a sample of $n = 32$ what is the sampling distribution of mussel shell lengths?

5. A botanist fits $\textit{Height} = \beta_0 + \beta_1 \cdot \textit{Light}$. Is this model phenomenological or mechanistic?

6. In a simple linear model, residuals ϵ_i are assumed i.i.d. $N(0, \sigma^2)$. What does $\mathbb{E}\epsilon_i = 0$ mean in words?

Group 2

7. A clutch has 3 eggs; let X = number that hatch with probabilities $P(X = 0) = 0.1,\ P(X = 1) = 0.25,\ P(X = 2) = 0.45,\ P(X = 3) = 0.2$. Compute $\mathbb{E}X$.

8. A field survey records daily sightings of a rare butterfly with pmf $P(Y = 0) = 0.5,\ P(Y = 1) = 0.3,\ P(Y = 2) = 0.15,\ P(Y = 3) = 0.05$. Find $\mathbb{V}(Y)$.

9. Wing lengths of a moth species are $N(6.0, 0.6)$ cm. What proportion exceeds 7.0 cm? (Use the z-table.)

10. Snail masses are $N(18, 3)$ g. Find $P(15 < m < 21)$.

11. Given $\mu = 28$ cm, $\sigma = 5$ cm and $n = 25$, compute $\mu_{\bar{X}}$ and $\sigma_{\bar{X}}$.

12. In a brood of $n = 4$ nestlings, each survives with probability $p = 0.6$. Let $X \sim \text{Bin}(n, p)$. Compute $E[X]$ and $\mathbb{V}(X)$.

13. Counts of a rare fern per $1\,\text{m}^2$ follow $X \sim \text{Pois}(1.2)$. Compute $\mathbb{E}X$.

14. A growth model $N(t) = N_0 e^{rt}$ with $N_0 = 50$ and $r = 0.12$. Find $N(5)$.

15. Given pairs (Sunlight hrs, Leaf area cm^2): $(2, 12), (4, 17), (6, 23)$. Fit the least-squares line $y = \beta_0 + \beta_1 x$ (show β_0, β_1 numerically).

16. Given the equation $\hat{y} = 10 + 0.8x$ for x = seed mass (mg) and y = plant height (cm), interpret the slope parameter.

Group 3

17. Let Z denote number of infected zooplankton in a net tow with pmf $z = 0, 1, 2, 3, 4$ and $P(Z = z) = \{0.3, 0.35, 0.2, 0.1, 0.05\}$. (a) Verify $\sum_z P(Z = z) = 1$. (b) Compute $E[Z]$. (c) Compute $\mathbb{V}(Z)$.

18. A discrete r.v. X has support $\{0, 1, 2, 3\}$ with $P(X = 0) = a,\ P(X = 1) = 2a,\ P(X = 2) = 3a,\ P(X = 3) = 2a$. Find $E[X]$.

19. Tree ring widths are $N(2.5, 0.4^2)$ mm. (a) Compute $P(W > 3.0)$. (b) Compute the 10th percentile.

20. Body temperatures of lizards are $N(34.0, 1.2^2)$ °C. (a) Find $P(33.0 < T < 36.0)$. (b) What T cuts off the top 2.5%?

21. Shell thickness has $\sigma = 0.08$ mm. For $n = 36$ oysters, find $P(\bar{X} > 1.05)$ when $\mu = 1.00$ mm.

22. Nectar sugar (%) is $N(22, 3)$. For $n = 25$, find $P(\bar{X} < 21)$.

23. Let $X \sim \text{Bin}(8, 0.55)$. (a) Compute $P(X = 3)$.

24. Daily arrivals of shorebirds at a cove follow $X \sim \text{Pois}(7)$. (a) $P(X = 4)$? (b) $P(X \leq 4)$? (c) $E[X]$ and $\mathbb{V}(X)$?

25. A regression predicts fish length (cm) from age (months): $\hat{y} = 5.2 + 0.9x$. (a) Predict length at $x = 8$. (b) If the observed $y = 12.0$, compute the residual.

26. A simple linear model for salamander mass (g) on tail length (cm) yields residuals that fan out with increasing x. (a) Which assumption is violated? (b) Name one solution to this violation.

Group 4

27. Let Y have pmf on $\{0, 1, 2, 3, 4, 5\}$ with $P = \{0.08, 0.22, 0.28, 0.22, 0.15, 0.05\}$. (a) Compute $E[Y]$. (b) Compute $\mathbb{V}(Y)$. (c) Compute $P(Y \geq 3)$.

28. A discrete r.v. X on $\{0, 1, 2, 3\}$ has frequencies $\{1, 3, 3, 1\}$. Compute $E[X]$ and $\mathbb{V}(X)$.

29. Stem diameters are $N(10, 1.8^2)$ mm. (a) Find $P(7.5 < D < 12.5)$. (b) Determine d with $P(D < d) = 0.90$.

30. Egg weights are $N(62, 6^2)$ g. (a) Compute the probability a random egg exceeds 70 g. (b) For a dozen eggs, approximate the probability the sample mean exceeds 65 g.

31. Leaf nitrogen has $\sigma = 0.9\%$. For $n = 49$, (a) compute $\sigma_{\bar{X}}$. (b) If $\mu = 3.0\%$, find $P(2.8 < \bar{X} < 3.2)$.

32. Fish mercury is $N(0.40, 0.12)$ ppm. For $n = 36$, find $P(\bar{X} > 0.45)$.

33. Given $X \sim \text{Beta}(1, 4)$. Find $\mathbb{E}X$ and $\mathbb{V}X$

34. Insect counts per trap follow $X \sim \text{Pois}(2.5)$. (a) $P(X \geq 4)$? (b) $P(2 \leq X \leq 5)$? (c) $E[X]$.

35. Regression of leaf area (y) on sunlight (x) gives $\hat{y} = 8.0 + 2.1x$ and $R^2 = 0.65$. (a) Interpret slope. (b) Predict at $x = 7$. (c) Interpret R^2.

36. A linear model $\hat{y} = 15 + 0.5x$ for shell mass on length yields residuals roughly normal with constant spread. (a) State the three core SLR assumptions checked here. (b) Which appear reasonable and why?

Group 5

37. Let X have pmf on $\{0, 1, \ldots, 6\}$ with $P(X = k) \propto \binom{6}{k}(0.4)^k(0.6)^{6-k}$. (a) What is the name of this distribution? (b) What is the formula for this distribution's expected value?

38. Given a set of two continuous variables (q, w) that show approximately bell shaped histograms when graphed, propose a model for determining the effect of w on q.

39. If the variables q and w from the previous question are determined to be proportions, would the assumptions of a linear model hold? Why or why not?

40. If the variables q and w are determined to be binary, and count data respectively, what distributions could be used to represent those variables? Propose a method for predicting the effect of q on w given this.

41. For $X \sim N(100, 15)$ (tree heights in cm), (a) Find a with $P(X < a) = 0.025$. (b) Find b with $P(X > b) = 0.10$. (c) Compute $P(a < X < b)$ using your a, b.

42. Fish lengths $L \sim N(30, 5^2)$. A regulation protects the middle 80% of the distribution. (a) Find the lower and upper cutoffs. (b) What fraction would be legal to harvest under this rule?

43. Leaf chlorophyll is $N(\mu = 42, \sigma = 6)$. For $n = 64$, (a) Compute $\sigma_{\bar{X}}$. (b) Find $P(41 < \bar{X} < 43)$. (c) If a single sample mean observed is $\bar{X} = 44.5$, compute its z relative to $\mu_{\bar{X}}$ and interpret.

44. Turtle mass has $\sigma = 220$ g. For $n = 49$, a study reports $\bar{X} = 1800$ g, $\mu = 1700$ g. (a) Compute $\sigma_{\bar{X}}$. (b) Compute $z = (\bar{X} - \mu)/\sigma_{\bar{X}}$. (c) Approximate $P(\bar{X} \geq 1800)$.

45. Consider $N(t) = N_0 e^{rt}$ with observation noise: $Y(t) = N(t) + \epsilon(t)$, $\epsilon(t) \overset{\text{i.i.d.}}{\sim} N(0, \sigma^2)$. (a) Identify deterministic vs probabilistic parts. (b) Is this linear or nonlinear in parameters? Explain. (c) Log-transform to obtain a linear relation in parameters and write the transformed model.

46. Least squares fit for seed mass (mg) on moisture (%) using data (x, y): $(5, 12), (10, 16), (15, 17), (20,$ (a) Compute $\hat{\beta}_1$ and $\hat{\beta}_0$. (b) Report $\hat{y}$ at $x = 18$. (c) Comment on whether the fit appears to under/over-predict at low x based on signs of residuals (briefly).

47. A regression output reports $R^2 = 0.88$ for predicting larval mass from food density. (a) Explain R^2 in this context. (b) Give two reasons why a relatively high R^2 does not guarantee causal interpretation.

48. A simple linear model for beetle mass on elytra length gives $\hat{y} = a + bx$ with residual plot showing a curved pattern. (a) Which assumption is likely violated? (b) Suggest a modeling remedy and write the new model. (c) Briefly describe how you would check if the remedy improved fit.

49. For $\hat{y} = 22+1.4x$ (plant biomass vs. light). (a) Predict $\hat{y}$ at $x = 12$. (b) An observation at $x = 12$ equals $y = 45$; compute the residual. (c) The residuals histogram is right-skewed: what SLR assumption might be questionable, and what is a possible fix?

50. A simple linear model assumes $\epsilon_i \sim N(0, \sigma^2)$. (a) State two practical consequences of this assumption for inference on β_1. (b) If residuals show increasing variance with $\hat{y}$, discuss how this affects standard errors and a potential modeling change.