STATSB-CA · Statistics B — Probability Models and Statistical Inference (California)

# Unit 1: Continuous Distributions and Sampling Distributions

Practice Workbook. The same problems, in the same order, as the Practice sections of the online lessons.

Name ______________________________ Period __________

## Lesson 1.1 — Continuous Random Variables and Density Curves

### FLUENCY

1. A wait W is uniformly distributed from 0 to 20 minutes. Find the height of the density, $P(W > 15)$, $P(5 < W < 8)$ and the mean wait.

2. For the same uniform wait, find $P(W = 7)$.

3. A waiting time is exponential with mean 5 minutes. Find $P(X > 5)$.

4. For the same exponential wait (mean 5), find $P(X < 2)$.

5. Find the median of an exponential distribution with mean 5.

6. An exponential wait has mean 10 minutes. Find $P(X > 20)$.

7. Can a density curve have height greater than 1 anywhere? Explain.

8. A density curve is symmetric about 12. What is $P(X < 12)$?

## APPLICATION

9. (Illustrative) Calls to a school's attendance line arrive at random, on average every 3 minutes. What is the probability of waiting more than 6 minutes for the next call? More than 9 minutes, given that 3 minutes have already passed?

10. A light-rail train arrives exactly every 15 minutes, and you arrive at a random time. Which model fits your wait, uniform or exponential? Find $P(\text{wait} > 10)$.

11. For an exponential wait with mean 4 minutes, find $P(2 < X < 6)$ and explain why the mean exceeds the median.

## CHALLENGE

12. The density $f(x) = cx$ on $0 \le x \le 2$ (and 0 elsewhere) is a valid density curve for one value of c. Find c, then find $P(X < 1)$.

## REVIEW

13. $X \sim \text{Binomial}(n = 15, p = 0.3)$. Find $P(X = 2)$, $P(X \le 2)$, and the mean and standard deviation of $X$.

14. A game pays –10 with probability 4/13; 5 with probability 5/13; 50 with probability 4/13. Find the expected value.

## Lesson 1.2 — The Normal Distribution

### FLUENCY

1. A distribution is approximately normal with mean 500 and standard deviation 100. Find the $z$-score of 562.3 and the proportion of values below it.

2. A distribution is approximately normal with mean 500 and standard deviation 100. Find the $z$-score of 704.0 and the proportion of values below it.

3. A distribution is approximately normal with mean 3.1 and standard deviation 0.4. Find the $z$-score of 3.8 and the proportion of values below it.

4. A distribution is approximately normal with mean 3.1 and standard deviation 0.4. Find the $z$-score of 3.2 and the proportion of values below it.

5. $X$ is normal with mean 100 and standard deviation 15. Find $P(88 < X < 105)$.

6. $X$ is normal with mean 100 and standard deviation 15. Find $P(83 < X < 109)$.

7. $X$ is normal with mean 70 and standard deviation 8. Find $P(55 < X < 77)$.

8. $X$ is normal with mean 500 and standard deviation 100. Find $P(459 < X < 592)$.

9. For a standard normal Z, find $P(-1.96 < Z < 1.96)$ and the 90th percentile of Z.

## APPLICATION

10. San Diego annual mean temperature, 1950-2024 (NOAA): mean 64.11°F, SD 1.42°F. Check the 68-95-99.7 rule: 65.3% of the 75 years are within 1 SD and 93.3% within 2 SD. Then use a normal model to estimate the fraction of years at 66°F or above, and compare with the data (10 years).

11. Using $N(64.11, 1.42)$, find the temperature that marks the warmest 10% of San Diego years, and compare with the data's 90th percentile (66.16°F).

12. A normal distribution has 25th percentile 62.0 and 75th percentile 64.4. Find $\mu$ and $\sigma$.

## CHALLENGE

13. A normal distribution has 10% of its values below 50 and 30% above 70. Find $\mu$ and $\sigma$.

## REVIEW

14. Use technology to find the least-squares line and the correlation coefficient for (3, 12), (4, 17), (6, 21), (7, 22), (12, 33), (13, 32).

15. A survey records two yes/no questions. Yes to both: 9; yes to A only: 47; yes to B only: 58; no to both: 45. Find P(A), P(A and B), P(A | B), and decide whether A and B look independent.

## Lesson 1.3 — When Is a Normal Model Appropriate?

### FLUENCY

1. $X \sim \text{Binomial}(n = 7, p = 0.4)$. Find $P(X = 5)$, $P(X \le 5)$, and the mean and standard deviation of $X$.

2. $X \sim \text{Binomial}(n = 10, p = 0.5)$. Find $P(X = 9)$, $P(X \le 9)$, and the mean and standard deviation of $X$.

3. Check the normal-approximation conditions for Binomial(40, 0.1).

4. Check the conditions for Binomial(200, 0.3).

5. Binomial(100, 0.5): find the mean and SD of the approximating normal distribution.

6. To approximate $P(X \le 12)$ for a binomial count with a normal curve, which boundary should you use?

7. To approximate $P(X \ge 25)$, which boundary should you use?

8. $X \sim$ Binomial(100, 0.1675) (rainy days in a random sample of 100 SFO days). Find $P(X \leq 12)$ exactly and with the normal approximation (continuity correction).

9. $X \sim$ Binomial(50, 0.4). Find $P(X \geq 25)$ exactly and by the normal approximation.

**APPLICATION**

10. SFO calendar-year precipitation, 1950-2024: mean 19.76 in, SD 6.79 in. A normal model gives $P(X < 8) \approx 0.042$; the data have 2 of 75 years below 8 inches. The mean (19.76) and median (19.54) are close. Is a normal model reasonable here? Compare with LAX rainfall.

11. Fresno annual mean temperature, 1950-2024: 65.3% of years within 1 SD and 97.3% within 2 SD of the mean. The record also has a strong upward trend (see the era comparison in Statistics A, Lesson 1.5). Is a single normal model a good description of these 75 values?

12. A student says: 'n = 30, so the binomial with p = 0.05 is approximately normal.' Evaluate the claim.

## CHALLENGE

13. For p = 0.2, what is the smallest n for which both $np \geq 10$ and $n(1 - p) \geq 10$? Explain why the binomial becomes more normal as n grows.

## REVIEW

14. (Statistics A) Why did the binomial model fail for rainy days per week at SFO?

15. (Statistics A) Find the z-score of 25.37 in a distribution with mean 12.11 and SD 5.69.

## Lesson 1.4 — Sampling Distributions and the Central Limit Theorem

### FLUENCY

1. Daily highs at SFO (2000-2024) have $\sigma \approx 8.30$°F. Find the SD of $\bar{x}$ for random samples of 25 days, 64 days and 100 days.

2. The population mean daily high is 66.30°F. What is the mean of the sampling distribution of $\bar{x}$ for $n = 40$?

3. The proportion of SFO days at or above 75°F is $p \approx 0.1384$. Find the SD of $\hat{p}$ for $n = 50$ and $n = 200$.

4. A population has $\sigma = 12$. What sample size makes the SD of $\bar{x}$ equal to 2?

5. Is the 10% condition met for a sample of 500 days from the 9,112-day SFO record?

6. By what factor does the SD of $\bar{x}$ change when n goes from 25 to 100?

7. True or false: the central limit theorem says that a large sample of data will be approximately normal.

## APPLICATION

8. Use the CLT to estimate the probability that a random sample of 36 SFO days has a mean high above 70°F ($\mu = 66.30$, $\sigma = 8.30$). A simulation of 5,000 such samples gave a proportion of 0.0050; compare.

9. A reporter takes the first 30 days of July at SFO as a 'sample' to estimate the station's mean annual daily high. Does the CLT make $\bar{x}$ a good estimate? Explain.

10. Daily rainfall at SFO ($\sigma \approx 0.207$ in, extremely skewed) and daily highs ($\sigma \approx 8.30$°F, mildly skewed): for which population is $n = 30$ enough for $\bar{x}$ to be approximately normal? Explain.

### CHALLENGE

11. What is the smallest sample size for which the SD of $\bar{x}$ for SFO daily highs is at most 1°F?

### REVIEW

12. $X$ is normal with mean 70 and standard deviation 8. Find $P(53 < X < 87)$.

13. (Lesson 1.2) For N(66.30, 8.30), find $P(X > 75)$.

14. (Statistics A) A random variable has $E(X) = 5$ and SD 2. Find the mean and SD of $3X + 1$.

## Unit 1 checkpoint

1. True or false: For a continuous random variable, $P(X = 3)$ is the height of the density at 3.
2. For a uniform wait from 0 to 12 minutes, P(wait > 9) = ______________.
3. For an exponential wait with mean 4, $P(X > 6) = e^{(-1.5)} \approx$ ______________.
4. A normal model with mean 63.15 and SD 1.27: about what fraction of values exceed 65?
   - A. 0.93
   - B. 0.072
   - C. 0.16
   - D. 0.5

5. A normal distribution with 10th percentile 61.5 and 90th percentile 64.8 has mean ______________.
6. Which data set is least suited to a normal model?
   A. Annual mean temperature at LAX
   B. Annual rainfall at LAX
   C. Heights of adult women
   D. Measurement errors
7. True or false: The normal approximation to Binomial(14, 0.5) for $P(X \leq 5)$ is much better with a continuity correction.
8. The SD of $\bar{x}$ for samples of 100 from a population with $\sigma = 8.3$ is ______________.
9. The mean of the sampling distribution of $\bar{x}$ equals…
   A. $\sigma/\sqrt{n}$
   B. $\mu$
   C. $\bar{x}$ from one sample
   D. 0
10. True or false: The central limit theorem makes a sample of 100 rainfall days approximately normal.
11. For a very skewed population, the sampling distribution of $\bar{x}$ at $n = 30$…
    A. Is always exactly normal
    B. May still be noticeably skewed
    C. Is uniform
    D. Has SD $\sigma$
12. True or false: A convenience sample's mean is unbiased if n is large enough.