PSTAT 5A Practice Worksheet 5 - SOLUTIONS

Continuous Random Variables and Confidence Intervals

Instructor Solutions

2025-07-29

Table of contents

1 Section A Solutions: Continuous Random Variables

Solution. **Solution A1: Distribution Identification and Properties**

(a) Exponential Distribution

Since the average time between arrivals is 2 minutes, we have:

- **Parameter λ:** The rate parameter $\lambda = \frac{1}{\mu} = \frac{1}{2} = 0.5$ arrivals per minute
- **Probability calculation:** $P(X \leq 1)$ where $X \sim Exponential(0.5)$

  For exponential distribution: $P(X \leq x) = 1 - e^{(-\lambda x)}$

  $P(X \leq 1) = 1 - e^{(-0.5 \times 1)} = 1 - e^{(-0.5)} = 1 - 0.6065 = \boxed{0.3935}$

(b) Uniform Distribution

- **Parameters:** $a = 10, b = 30$
- **Expected Value:** $E[X] = (a+b)/2 = \frac{(10+30)}{2} = \boxed{20}$
- **Variance:** $Var(X) = \frac{(b-a)^2}{12} = \frac{(30-10)^2}{12} = \frac{400}{12} = \boxed{33.3333}$

Solution. **Solution A2: Normal Distribution Calculations**

Given: $X \sim N(64, 2.5^2)$

(a) $P(X > 67)$

Step 1: Standardize

$Z = (67 - 64)/2.5 = 3/2.5 = 1.2$

Step 2: Find probability

$P(X > 67) = P(Z > 1.2) = 1 - P(Z \leq 1.2) = 1 - 0.8849 = \boxed{0.1151}$

(b) 25th percentile

Step 1: Find z-value for 25 -th percentile

$P(Z \leq z) = 0.25$, so $z_{0.25} = -0.6745$

Step 2: Convert back to X

$x = \mu + z\sigma = 64 + (-0.6745)(2.5) = 64 - 1.6863 = \boxed{62.3137}$ inches

(c) P(62 < X < 68)

Step 1: Standardize both values

$Z_1 = (62 - 64)/2.5 = -0.8$ $Z_2 = (68 - 64)/2.5 = 1.6$

Step 2: Find probability

$P(62 < X < 68) = P(-0.8 < Z < 1.6) = P(Z < 1.6) - P(Z < -0.8)$

$= 0.9452 - 0.2119 = \boxed{0.7333}$

2 Section B Solutions: Confidence Intervals

Solution. **Solution B1: Understanding Confidence Intervals**

(a) Explanation of 95% Confidence Interval:

A 95% confidence interval means that if we were to repeat our sampling process many times (say 100 times) and construct a confidence interval each time using the same method, approximately 95 of those intervals would contain the true population mean. **It does NOT mean there's a 95% probability that the population mean lies in any one specific interval.**

(b) Sample mean and margin of error:

Given CI: $(150g, 170g)$

- **Sample mean:** $\bar{x} = (150 + 170)/2 = \boxed{160g}$
- **Margin of error:** $E = (170 - 150)/2 = \boxed{10g}$

(c) True or False statement:

FALSE. Once we calculate a specific confidence interval, the population mean either is or isn't in that interval, there's no probability involved for that specific interval. The 95% refers to the long-run success rate of the method, not the probability for any individual interval.

Solution. **Solution B2: Constructing Confidence Intervals**

Given: $n = 36, \bar{x} = 78.5, s = 12$

(a) 95% Confidence Interval:

Step 1: Check conditions

- $n = 36 \geq 30$, so we can use z-distribution
- For 95

Step 2: Calculate margin of error

$E = z_{0.025} \times (\frac{s}{\sqrt{n}}) = 1.96 \times (\frac{12}{\sqrt{36}}) = 1.96 \times (\frac{12}{6}) = 1.96 \times 2 = 3.92$

Step 3: Construct interval

$CI = \bar{x} \pm E = 78.5 \pm 3.92 = \boxed{(74.58, 82.42)}$

(b) Interpretation:

We are 95% confident that the true population mean test score is between 74.58 and 82.42 points.

(c) Effects on interval width:

- **Increasing confidence level to 99%:** The interval would become **wider** because we need $z_{0.005} = 2.576 > 1.96$
- **Increasing sample size to 144:** The interval would become **narrower** because the margin of error would be $E = 1.96 \times (\frac{12}{\sqrt{144}}) = 1.96 \times 1 = 1.96$ (smaller than 3.92)

Solution. **Solution B3: Sample Size Determination**

Given: $E = \$5$, confidence $= 95\%, \sigma = \$25$

(a) Required sample size:

Step 1: Use sample size formula

$n = (z_{0.025} \times \frac{\sigma}{E})^2$

Step 2: Substitute values

$n = (1.96 \times 25/5)^2 = (49/5)^2 = 9.8^2 = 96.04$

Step 3: Round up

$\boxed{n = 97}$ customers (**always round up for sample size**)

(b) For margin of error = \$3:

$n = (1.96 \times 25/3)^2 = (49/3)^2 = 16.333^2 = 266.67$

$\boxed{n = 267}$ customers

3 Optional Problem Solutions

Solution. **Optional Solution 1: Conceptual Understanding**

(a) Differences between discrete and continuous:

Values they can take:

- Discrete: Countable values (integers, specific points)
- Continuous: Uncountably infinite values (any real number in an interval)

How we calculate probabilities:

- Discrete: $P(X = x)$ can be non-zero; we sum probabilities

- Continuous: $P(X = x) = 0$ for any specific x; we integrate over intervals

(b) Why $P(X = x) = 0$ for continuous distributions:

In continuous distributions, there are infinitely many possible values in any interval. The probability of hitting any one exact value is infinitesimally small, hence zero. We instead calculate $P(a < X < b)$ by integrating the PDF over the interval $[a, b]$.

(c) Relationship between PDF and CDF:

- **PDF (f(x)):** The probability density function gives the "density" of probability at each point
- **CDF (F(x)):** The cumulative distribution function gives $P(X \leq x)$
- **Relationship:** $F(x) = \int_{-\infty}^{x} f(t)dt, \quad \text{and} \quad f(x) = F'(x)$

Key Takeaways:

1. **Always standardize** normal distribution problems using $Z = \frac{(X-\mu)}{\sigma}$
2. **Interpret confidence intervals** in context, they're about the method's reliability, not individual interval probabilities
3. **Choose the right distribution** use t when σ is unknown and $n < 30$
4. **Round up sample sizes** to ensure you meet the margin of error requirement
5. **For continuous distributions**, focus on intervals, not individual points