Chapter 9

Interval Estimation

Detailed Solutions

9.1 Basic Concept

9.1 True or False: The Meaning of Confidence

Problem: 95% CI is $[10, 20]$. Evaluate statements.
Solution:

(a) **False.** The population mean μ is a fixed constant, not a random variable. It is either in the interval or it is not. We cannot assign a probability to a fixed event after the fact.

(b) **True.** This is the correct frequentist interpretation. "Confidence" refers to the reliability of the method (the interval construction process), not the specific result of one experiment.

(c) **False.** A Confidence Interval estimates the *population parameter* (mean), not the distribution of individual data points. (That would be a Tolerance Interval).

(d) **True.** Before data collection, the interval boundaries $(\bar{X} \pm E)$ are random variables. After collection, they become fixed numbers $[10, 20]$.

$$\boxed{\text{(a) False} \quad \text{(b) True} \quad \text{(c) False} \quad \text{(d) True}}$$

9.2 Factors Affecting Width

Problem: Width $W = 2 \times z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$.
Solution:

(a) **Increase n (10 to 100): Narrower**. n is in the denominator $(\sqrt{n})$. More data reduces uncertainty.

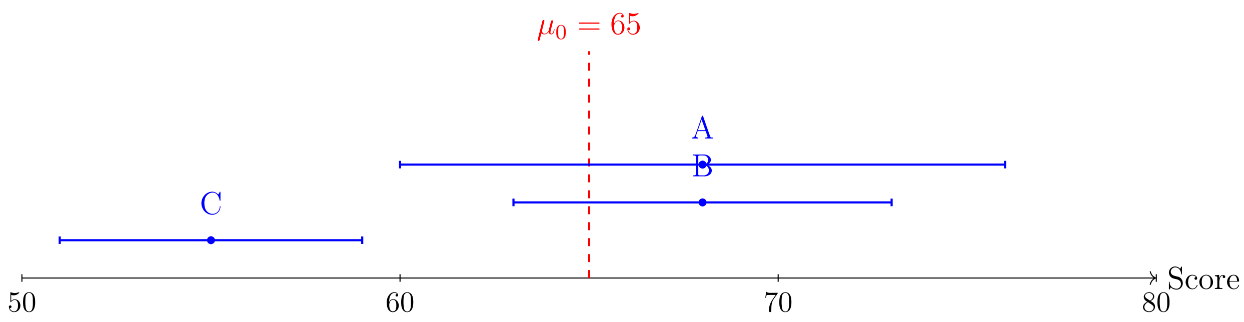
(b) **Increase Confidence (90% to 99%): Wider**. To be more confident that we capture the true mean, we must cast a wider net (z-score increases).

(c) **Larger s: Wider**. More variability in the data (noise) makes precise estimation harder.

(a) Narrower (b) Wider (c) Wider

9.3 Visual Forensics

Problem: Interpreting CIs from the chart.



Solution:

(a) **Sample Size:** Study B likely used a **larger sample size**. Both have the same center and assumed same variance, but Interval B ($73 - 63 = 10$) is narrower than Interval A ($76 - 60 = 16$). Since Width $\propto 1/\sqrt{n}$, a narrower width implies larger n.

(b) **Significance ($\mu_0 = 70$):** Interval A is $[60, 76]$. Since 70 is **inside** the interval, the mean is **not significantly different** from 70 at the 5% level. We cannot reject the null hypothesis that $\mu = 70$.

(c) **Overlap:** Interval B $[63, 73]$ and Interval C $[51, 59]$ do **not overlap**. This strongly suggests that the population means are significantly different.

(d) **Precision:** Study C has the narrowest interval ($59 - 51 = 8$). Thus, Study C provides the most precise estimate.

(a) Study B (Narrower)
(b) Not significant (70 is inside)
(c) No overlap
(d) Study C (Narrowest width)

9.2 Intermediate

9.4 CI for Mean (Raw Data)

Problem: Data: $\{52, 48, 56, 45, 50, 53, 49, 51, 54, 42\}$. $n = 10$.
Solution:

(a) **Stats:** $\sum x = 500 \implies \bar{x} = 50.0$. $s^2 = \frac{\sum(x_i - \bar{x})^2}{n-1} = \frac{148}{9} \approx 16.44$. $s = \sqrt{16.44} \approx 4.055$.

(b) **95% CI:** Since $n < 30$ and σ unknown, use t-distribution with $df = 9$. $t_{0.025,9} = 2.262$.

$$\begin{aligned} \bar{x} \pm t\frac{s}{\sqrt{n}} &= 50.0 \pm 2.262\,\frac{4.055}{\sqrt{10}} \\ &= 50.0 \pm 2.262(1.282) \\ &= 50.0 \pm 2.90. \end{aligned}$$

$$\text{CI} = [47.1,\ 52.9]\text{ MPa}.$$

(c) **Interpretation:** The standard requires $\mu \geq 48$. The interval is $[47.1, 52.9]$. Since the lower bound (47.1) is **less than** 48, we **cannot** guarantee with 95% confidence that the alloy meets the standard. It implies the mean *could* be as low as 47.1.

$$\boxed{\begin{array}{l} \text{(a) } \bar{x} = 50, s = 4.06 \\ \text{(b) } [47.1, 52.9] \text{ MPa} \\ \text{(c) No strong evidence (Lower bound} < 48\text{).} \end{array}}$$

9.5 Difference of Means

Problem: Catalyst A vs B. Equal variances assumed.
Solution:

(a) **Stats:**

A: $\{85, 88, 84, 86, 90, 83, 87, 85\}$. $n_A = 8, \bar{x}_A = 86.0$. $s_A^2 = \frac{1}{7}\sum(x-86)^2 = \frac{32}{7} \approx 4.57$.

B: $\{81, 78, 83, 82, 80, 79, 84, 81\}$. $n_B = 8, \bar{x}_B = 81.0$. $s_B^2 = \frac{1}{7}\sum(x-81)^2 = \frac{28}{7} = 4.00$.

(b) **Pooled Variance:**

$$s_p^2 = \frac{(n_A - 1)s_A^2 + (n_B - 1)s_B^2}{n_A + n_B - 2} = \frac{7(4.57) + 7(4.00)}{14} = \frac{32 + 28}{14} = \frac{60}{14} \approx 4.286$$

$s_p = \sqrt{4.286} \approx 2.07$.

(c) **95% CI:** $df = 14$. $t_{0.025,14} = 2.145$.

$$
\begin{aligned}
(\bar{x}_A - \bar{x}_B) \pm t \cdot s_p\sqrt{\frac{1}{n_A} + \frac{1}{n_B}} &= (86 - 81) \pm 2.145(2.07)\sqrt{\frac{1}{8} + \frac{1}{8}} \\
&= 5 \pm 2.145(2.07)(0.5) \\
&= 5 \pm 2.22.
\end{aligned}
$$

$$\text{CI} = [2.78,\ 7.22].$$

(d) **Conclusion:** The interval lies entirely above 0 (Lower bound $2.78 > 0$). Yes, Catalyst A produces a significantly higher yield.

(b) $s_p^2 \approx 4.29$ (c) $[2.78, 7.22]$ (d) Yes, significantly higher.

9.6 Variance Estimation

Problem: Data deviations: $\{-2, 1, 0, -1, 2, -3, 0, 1, -2, 1, 3, -1\}$. $n = 12$.
Solution:

(a) **Sample Variance:** $\sum x = -1$. $\bar{x} = -1/12 \approx -0.083$. $s^2 = \frac{\sum x^2 - (\sum x)^2/n}{n-1} = \frac{35-(-1)^2/12}{11} = \frac{34.917}{11} \approx 3.174$.

(b) **90% CI for σ^2:** χ^2 distribution with $df = 11$. $\alpha = 0.10$. $\chi^2_{0.05,11} = 19.675$ (Upper), $\chi^2_{0.95,11} = 4.575$ (Lower).

$$
\begin{aligned}
\left[\frac{(n-1)s^2}{\chi^2_{\text{upper}}}, \frac{(n-1)s^2}{\chi^2_{\text{lower}}}\right] &= \left[\frac{11(3.174)}{19.675}, \frac{11(3.174)}{4.575}\right] \\
&= \left[\frac{34.914}{19.675}, \frac{34.914}{4.575}\right] \\
&= [1.77,\ 7.63].
\end{aligned}
$$

(c) **Within Spec ($\sigma^2 \leq 4$)?** The interval extends up to 7.63. Since values > 4 are plausible (inside the CI), we **cannot** confirm with 90% confidence that the machine is within spec.

(a) $s^2 \approx 3.17$ (b) $[1.77, 7.63]$ (c) No, upper bound $7.63 > 4$.

9.7 Proportions (A/B Testing)

Problem: Red (50/1000) vs Blue (70/1000).
Solution:

(a) **Proportions:** $\hat{p}_R = 0.050$, $\hat{p}_B = 0.070$.

(b) **99% CI for Diff:**

$z_{0.005} = 2.576$.

$SE = \sqrt{\frac{p_R q_R}{n} + \frac{p_B q_B}{n}} = \sqrt{\frac{0.05(0.95)}{1000} + \frac{0.07(0.93)}{1000}} \approx 0.0106$.

Diff $= 0.070 - 0.050 = 0.020$.

$$CI = 0.020 \pm 2.576(0.0106) = 0.020 \pm 0.027$$

$$[-0.007, 0.047]$$

(c) **Conclusion:** The interval **contains 0**. This means the difference is not statistically significant at 99% confidence. We cannot definitively say Blue is better.

(b) $[-0.007, 0.047]$
(c) Contains 0. Difference not significant.

9.3 Challenge

9.8 Pivotal Quantity Method

Problem: Derive T-interval.
Solution:

(a) **Pivot:** $T = \frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t_{n-1}$. This is pivotal because its distribution (t-dist) is parameter-free.

(b) **Probability Statement:**

$$P(-t_{\alpha/2} < \frac{\bar{X}-\mu}{S/\sqrt{n}} < t_{\alpha/2}) = 1-\alpha$$

(c) **Rearrange:** Multiply by $S/\sqrt{n}$:

$$-t_{\alpha/2}\frac{S}{\sqrt{n}} < \bar{X}-\mu < t_{\alpha/2}\frac{S}{\sqrt{n}}$$

Subtract $\bar{X}$:

$$-\bar{X} - t_{\alpha/2}\frac{S}{\sqrt{n}} < -\mu < -\bar{X} + t_{\alpha/2}\frac{S}{\sqrt{n}}$$

Multiply by -1 (flip inequalities):

$$\bar{X} + t_{\alpha/2}\frac{S}{\sqrt{n}} > \mu > \bar{X} - t_{\alpha/2}\frac{S}{\sqrt{n}}$$

Rewrite:

$$\bar{X} - t_{\alpha/2}\frac{S}{\sqrt{n}} < \mu < \bar{X} + t_{\alpha/2}\frac{S}{\sqrt{n}}$$

$\boxed{\text{Proven.}}$

9.9 CI for Gamma Parameter

Problem: $Y = \frac{2}{\beta}\sum X_i \sim \chi^2_{2n\alpha}$.
Solution:

(a) **Validity:** Y is a function of data ($\sum X_i$) and parameter (β), and its distribution (χ^2) is known and fixed (independent of β). Thus, it is a pivot.

(b) **Derive:**

$$P\left(\chi^2_{\text{lower}} < \frac{2\sum X_i}{\beta} < \chi^2_{\text{upper}}\right) = 1-\alpha.$$

Invert the inequalities:

$$\frac{1}{\chi^2_{\text{lower}}} > \frac{\beta}{2\sum X_i} > \frac{1}{\chi^2_{\text{upper}}}.$$

Multiply through by $2\sum X_i$:

$$\frac{2\sum X_i}{\chi^2_{\text{upper}}} < \beta < \frac{2\sum X_i}{\chi^2_{\text{lower}}}.$$

$$\text{where } \chi^2_{\text{upper}} = \chi^2_{\alpha/2} \quad \text{and} \quad \chi^2_{\text{lower}} = \chi^2_{1-\alpha/2}.$$

(c) **Application:** Exponential $\alpha = 1, n = 10 \implies df = 2(10)(1) = 20$. $\sum x = 500$. $\chi^2_{0.025,20} = 34.17$, $\chi^2_{0.975,20} = 9.59$.

$$CI = \left[\frac{2(500)}{34.17}, \frac{2(500)}{9.59}\right] = \left[\frac{1000}{34.17}, \frac{1000}{9.59}\right] = [29.27, 104.28]$$

(b) $\left[\frac{2\sum X_i}{\chi^2_{\alpha/2}}, \frac{2\sum X_i}{\chi^2_{1-\alpha/2}}\right]$
(c) $[29.27, 104.28]$ hours

9.10 Prediction Interval vs CI

Problem: Error $X_{n+1} - \bar{X}$.
Solution:

(a) **Variance:** Since X_{n+1} is independent of the sample mean $\bar{X}$:

$$Var(X_{n+1} - \bar{X}) = Var(X_{n+1}) + Var(\bar{X}) = \sigma^2 + \frac{\sigma^2}{n} = \sigma^2\left(1 + \frac{1}{n}\right)$$

(b) **Formula:** The standard error is $s\sqrt{1 + 1/n}$.

$$PI = \bar{X} \pm t_{\alpha/2,n-1}s\sqrt{1 + \frac{1}{n}}$$

(Note: This interval is always wider than the CI because it accounts for the uncertainty of the mean PLUS the variation of the single new data point).

(a) $\sigma^2(1 + 1/n)$
(b) $\bar{X} \pm t \cdot s\sqrt{1 + 1/n}$

9.4 Application

9.11 Visualizing "Confidence"

Problem: Simulation of 100 Intervals.
Solution:

```
import numpy as np
import matplotlib.pyplot as plt
from scipy.stats import t

mu = 50
sigma = 10
n = 30
num_intervals = 100
confidence = 0.95

# Initialize
captured_count = 0
plt.figure(figsize=(10, 8))

for i in range(num_intervals):
    # 1. Generate Sample
    sample = np.random.normal(mu, sigma, n)
    x_bar = np.mean(sample)
    s = np.std(sample, ddof=1)

    # 2. Calculate CI
    crit_val = t.ppf((1 + confidence) / 2, df=n-1)
    margin_error = crit_val * (s / np.sqrt(n))

    low = x_bar - margin_error
    high = x_bar + margin_error

    # 3. Check Capture
    is_captured = low <= mu <= high
    if is_captured:
        captured_count += 1
        color = 'green'
    else:
        color = 'red'

    # 4. Plot
    plt.plot([low, high], [i, i], color=color, alpha=0.7)

# Finalize Plot
plt.axvline(mu, color='black', linestyle='--', linewidth=2, label='True Mean')
plt.title(f'100 Confidence Intervals Simulation\nCaptured: {captured_count}/100')
plt.xlabel('Value')
plt.ylabel('Simulation Index')
plt.legend()
plt.show()

# Interpretation:
# You should see approximately 95 green lines (captured) and 5 red
```

```
    lines (missed).
# This demonstrates that "95% Confidence" describes the long-run
    success rate of the method.
```