Random Intervals and Confidence Intervals (Ch. 6.1)

Yifan Zhu

Iowa State University

Outline

Motivation

Random Intervals

Confidence Intervals ($n \geq 25,\ \sigma$ known)

Statistical inference

- **Statistical inference**: using data from the sample to draw formal conclusions about the population
  - Point estimation (confidence intervals): estimating population parameters and specifying the degree of precision of the estimate.
  - Hypothesis testing: testing the validity of statements about the population that are framed in terms of parameters.

Motivation for confidence intervals

- We want information on a population. For example:
  - True mean breaking strength of a kind of wire rope.
  - True mean fill weight of food jars.
  - True mean instrumental drift of a kind of scale.
  - Average number of cycles to failure of a kind of spring.
- We can use point estimates:
  - For example: if we measure breaking strengths (in tons) of 6 wire ropes as 5, 3, 7, 3,10, and 1, we might estimate the true mean breaking strength $\mu \approx \overline{x} = \frac{5+3+7+3+10+1}{6} = 4.83$ tons.
- Or, we can use interval estimates:
  - μ is likely to be inside the interval $(4.83 - 2, 4.83 + 2) = (2.83, 6.83)$.
  - We are confident that the true mean breaking strength, μ, is somewhere in $(2.83, 6.83)$. But how confident can we be?

Outline

Motivation

Random Intervals

Confidence Intervals ($n \geq 25,\ \sigma$ known)

Random intervals

- A **random interval** is an interval on the real line with a random variable at one or both of the endpoints.
- Examples:
  - $(Z-2, Z+2)$, $Z \sim N(0,1)$
  - (Z, ∞)
  - $(-\infty, X)$, $X \sim N(-2, 9)$
  - $(T - s \cdot t_{7,0.975}, T + s \cdot t_{7,0.975})$, $T \sim t_7$
  - $(X - \sigma \cdot z_{1-\alpha}, \infty)$, $X \sim N(5, \sigma^2)$, $0 < \alpha < 1$.
- Random intervals take into account the uncertainty in the measurement of a true mean, μ.

Example: instrumental drift

- Let Z be a measure of instrumental drift of a random voltmeter that comes out of a certain factory. Say $Z \sim N(0, 1)$.
- Define a random interval:

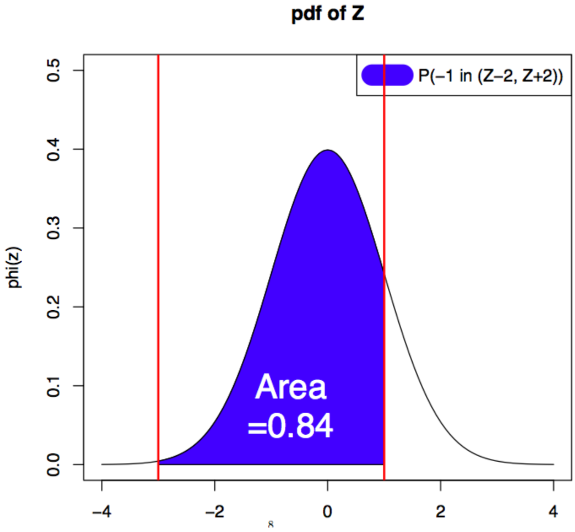
$$(Z - 2, Z + 2)$$

- What is the probability that -1 is inside the interval?
  - Equivalent to asking how likely it is that the drift of the next instrument is within 2 units of -1.

Example: instrumental drift

$$
\begin{aligned}
P(-1 \text{ in } (Z-2, Z+2)) &= P(Z-2 < -1 < Z+2) \\
&= P(Z-1 < 0 < Z+3) \\
&= P(-1 < -Z < 3) \\
&= P(-3 < Z < 1) \\
&= P(Z \leq 1) - P(Z \leq -3) \\
&= \Phi(1) - \Phi(-3) \\
&= 0.84
\end{aligned}
$$

Example: instrumental drift: the range of Z values for which -1 is in $(Z-2, Z+2)$



Your turn: random intervals

Calculate:

1. $P(2 \text{ in } (X-1, X+1)),\ X \sim N(2,4)$
2. $P(6.6 \text{ in } (X-2, X+1)),\ X \sim N(7,2)$

Here, $0 < \alpha < 1$.

Answers: random intervals

1. $X \sim N(2, 4)$

$$\begin{aligned}
P(2 \in (X-1, X+1)) &= P(X - 1 < 2 < X + 1) \\
&= P(-1 < 2 - X < 1) \\
&= P(-1 < X - 2 < 1) \\
&= P\left(\frac{-1}{2} < \frac{X-2}{2} < \frac{1}{2}\right) \\
&= P(-0.5 < Z < 0.5) \\
&= \Phi(0.5) - \Phi(-0.5) \\
&= 0.69 - 0.31 \\
&= 0.38
\end{aligned}$$

Answers: random intervals

2. $X \sim N(7, 2)$

$$
\begin{aligned}
P(6.6 \in (X-2, X+1)) &= P(X - 2 < 6.6 < X + 1) \\
&= P(-2 < 6.6 - X < 1) \\
&= P(-1 < X - 6.6 < 2) \\
&= P(-1.4 < X - 7 < 1.6) \\
&= P\left(\frac{-1.4}{\sqrt{2}} < \frac{X-7}{\sqrt{2}} < \frac{1.6}{\sqrt{2}}\right) \\
&= P(-0.99 < Z < 1.13) \\
&= \Phi(1.13) - \Phi(-0.99) \\
&= 0.87 - 0.16 \\
&= 0.71
\end{aligned}
$$

More abstract random intervals

- Let's say $X_1, X_2, \ldots, X_n$ are iid with:
  - $n \geq 25$
  - mean μ
  - variance σ^2
- The random interval, $(\overline{X} - z_{1-\alpha}\frac{\sigma}{\sqrt{n}},\ \infty)$, is useful for estimating μ $(0 < \alpha < 1)$.
- The interval contains μ with probability $1 - \alpha$.

$$
\begin{aligned}
&P(\mu \in (\overline{X} - z_{1-\alpha}\frac{\sigma}{\sqrt{n}},\ \infty)) \\
&\qquad = P\left(\overline{X} - z_{1-\alpha}\frac{\sigma}{\sqrt{n}} < \mu\right) \\
&\qquad = P\left(\overline{X} - \mu < z_{1-\alpha}\frac{\sigma}{\sqrt{n}}\right) \\
&\qquad = P\left(\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} < z_{1-\alpha}\right) \\
&\qquad \approx P(Z < z_{1-\alpha}) \quad \text{(Central Limit Theorem)} \\
&\qquad = \Phi(z_{1-\alpha}) \\
&\qquad = 1 - \alpha \quad \text{(by the definition of } z_p\text{)}
\end{aligned}
$$

Your turn: abstract random intervals

Calculate:

1. $P(\mu \in (-\infty,\ \overline{X} + z_{1-\alpha}\frac{\sigma}{\sqrt{n}})),\ \overline{X} \sim N(\mu, \sigma^2)$
2. $P(\mu \in (\overline{X} - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}},\ \overline{X} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}})),\ X \sim N(\mu, \sigma^2)$

Remember the Central Limit Theorem:

$$\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \approx N(0, 1)$$

Answers: abstract random intervals

1.

$$
\begin{aligned}
P(\mu \in & (-\infty,\ \overline{X} + z_{1-\alpha}\frac{\sigma}{\sqrt{n}})) \\
&= P\left(\mu < \overline{X} + z_{1-\alpha}\frac{\sigma}{\sqrt{n}}\right) \\
&= P\left(-z_{1-\alpha}\frac{\sigma}{\sqrt{n}} < \overline{X} - \mu\right) \\
&= P\left(-z_{1-\alpha} < \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}\right) \\
&\approx P\left(-z_{1-\alpha} < Z\right) \quad \text{(Central Limit Theorem)} \\
&= 1 - P(Z \leq -z_{1-\alpha}) \\
&= 1 - \Phi(-z_{1-\alpha}) \\
&= 1 - \Phi(z_{\alpha}) \quad \text{(by symmetry: N(0,1) pdf)} \\
&= 1 - \alpha \quad \text{(by the definition of } z_p\text{)}
\end{aligned}
$$

Answers: abstract random intervals

2.

$$
\begin{aligned}
P(\mu \in (X - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}},\ \overline{X} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}))& \\
&= P\left(\overline{X} - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \overline{X} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right) \\
&= P\left(-z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} < \mu - \overline{X} < z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right) \\
&= P\left(-z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} < \overline{X} - \mu < z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right) \\
&= P\left(-z_{1-\alpha/2} < \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} < z_{1-\alpha/2}\right) \\
&\approx P(-z_{1-\alpha/2} < Z < z_{1-\alpha/2}) \quad \text{(Central Limit Theorem)} \\
&= \Phi(z_{1-\alpha/2}) - \Phi(-z_{1-\alpha/2}) \\
&= \Phi(z_{1-\alpha/2}) - \Phi(z_{\alpha/2}) \quad \text{(by symmetry: N(0,1) pdf)} \\
&= (1 - \frac{\alpha}{2}) - \frac{\alpha}{2} = 1 - \alpha
\end{aligned}
$$

Outline

Motivation

Random Intervals

Confidence Intervals ($n \geq 25$, σ known)

Confidence intervals

- A $1-\alpha$ **confidence interval** for an unknown parameter is the finite realization of a random interval that contains that parameter with probability $1-\alpha$.
- $1-\alpha$ is called the **confidence level** of the interval.
- Example: for observations $x_1, x_2, \dots x_n$ from random variables $X_1, X_2, \dots, X_n$ iid with $E(X_1) = \mu$, $Var(X_1) = \sigma^2$, a $1-\alpha$ confidence interval for μ is:

$$\left(\overline{x} - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}, \overline{x} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$$

which is a random draw from the random interval:

$$\left(\overline{X} - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}, \overline{X} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$$

Confidence intervals for μ: σ known, $n \geq 25$

- Two-sided $1-\alpha$ confidence interval:

$$\left(\overline{x} - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}, \overline{x} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$$

- One-sided $1-\alpha$ **upper confidence interval**:

$$\left(-\infty,\ \overline{x} + z_{1-\alpha}\frac{\sigma}{\sqrt{n}}\right)$$

- One-sided $1-\alpha$ **lower confidence interval**:

$$\left(\overline{x} - z_{1-\alpha}\frac{\sigma}{\sqrt{n}},\ \infty\right)$$

Example: fill weight of jars

- Suppose a manufacturer fills jars of food using a stable filling process with a known standard deviation of $\sigma = 1.6g$.
- We take a sample of n =47 jars and measure the sample mean weight $\overline{x} = 138.2$ g.
- A two-sided 90% confidence interval ($\alpha = 0.1$) for the true mean weight μ is:

$$\left(\overline{x} - z_{1-0.1/2}\frac{\sigma}{\sqrt{n}},\ \overline{x} + z_{1-0.1/2}\frac{\sigma}{\sqrt{n}}\right)$$
$$= \left(138.2 - z_{0.95}\frac{1.6}{\sqrt{47}},\ 138.2 + z_{0.95}\frac{1.6}{\sqrt{47}}\right)$$
$$= (138.2 - 1.64 \cdot 0.23,\ 138.2 + 1.64 \cdot 0.23)$$
$$= (137.82, 138.58)$$

I could have also written the interval as:

$$138.2 \pm 0.38\ g$$

Interpreting the confidence interval: fill weight of jars

- We are 90% confident that the true mean fill weight is between 137.82g and 138.58g.
- If we took 100 more samples of 47 jars each, roughly 90 of those samples would yield confidence intervals containing the true mean fill weight.
- These methods of interpretation generalize to all confidence intervals.

Example: fill weight of jars.

- What if we just want to be sure that the true mean fill weight is high enough?
- Then, we would use a one-side lower 90% confidence interval:

$$
\begin{aligned}
&\left(\overline{x} - z_{1-\alpha}\frac{\sigma}{\sqrt{n}},\ \infty\right) \\
&= \left(138.2 - z_{1-\alpha}\frac{\sigma}{\sqrt{n}},\ \infty\right) \\
&= \left(138.2 - z_{0.9}\frac{1.6}{\sqrt{47}},\ \infty\right) \\
&= (138.2 - 1.28 \cdot 0.23,\ \infty) \\
&= (137.91,\ \infty)
\end{aligned}
$$

- We're 90% confident that the true mean fill weight is above 137.91 g.

Your turn: car engines

- Consider a grinding process used to rebuild car engines, which involves grinding rod journals for engine crankshafts.
- Of interest is the deviation of the true mean rod journal diameter from the target diameter.
- Suppose the standard deviation of the individual differences from the target diameter is 0.7×10^{-4} in.
- 32 consecutive rod journals are ground, with a sample mean deviation of -0.16×10^{-4} in from the target diameter.
- Calculate and interpret a two-sided 95% confidence interval for the true mean deviation from the target diameter. Is there enough evidence that we're missing the target on average?

Answer: car engines

- $\alpha = 1 - 0.95 = 0.05$, $n = 32$, $\sigma = 0.7 \times 10^{-4}$, and $\overline{x} = -0.16 \times 10^{-4}$.
- Interval:

$$\left(\overline{x} - z_{1-0.05/2}\frac{\sigma}{\sqrt{n}},\ \overline{x} + z_{1-0.05/2}\frac{\sigma}{\sqrt{n}}\right)$$
$$= \left(-0.16 \times 10^{-4} - z_{0.975}\frac{0.7 \times 10^{-4}}{\sqrt{32}},\ -0.16 \times 10^{-4} + z_{0.975}\frac{0.7 \times 10^{-4}}{\sqrt{32}}\right)$$
$$= (-0.16 \times 10^{-4} - 1.96 \cdot 1.2 \times 10^{-5},\ -0.16 \times 10^{-4} + 1.96 \cdot 1.2 \times 10^{-5})$$
$$= (-4.0 \times 10^{-5}, 7.5 \times 10^{-6})$$

- We are 95% confident that the true mean deviation from the target diameter of the rod journals is between -4.0×10^{-5} in and 7.5×10^{-6} in.
- Since 0 is in the confidence interval, there is not enough evidence to conclude that the rod journal grinding process is off target.

Your turn: hard disk failures

- F. Willett, in the article *The Case of the Derailed Disk Drives* (Mechanical Engineering, 1988), discusses a study done to isolate the cause of *blink code A failure* in a model of Winchester hard disk drive.
- For each disk, the investigator measured the breakaway torque (in. oz.) required to loosen the drive's interrupter flag on the stepper motor shaft.
- Breakaway torques for 26 disk drives were recorded, with a sample mean of 11.5 in. oz.
- Suppose you know the true standard deviation of the breakaway torques is 5.1 in. oz.
- Calculate and interpret:
  1. A two-sided 90% confidence interval for the true mean breakaway torque of the relevant type of Winchester drive.
  2. An analogous two-sided 95% confidence interval.
  3. An analogous two-sided 99% confidence interval.
- Is there enough evidence to conclude that the mean breakaway torque is different from the factory's standard of 33.5 in. oz.?

Answers: hard disk failures

- $\sigma = 5.1, \overline{x} = 11.5, n = 26.$
- All three confidence intervals have the form:

$$\left(\overline{x} - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}},\ \overline{x} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$$
$$= \left(11.5 - z_{1-\alpha/2}\frac{5.1}{\sqrt{26}},\ 11.5 + z_{1-\alpha/2}\frac{5.1}{\sqrt{26}}\right)$$
$$= (11.5 - 1.0002 \cdot z_{1-\alpha/2},\ 11.5 + 1.0002 \cdot z_{1-\alpha/2})$$

- The confidence intervals are thus:
  1. 90% CI means $\alpha = 0.1$

$$(11.5 - 1.0002 \cdot z_{1-0.1/2},\ 11.5 + 1.0002 \cdot z_{1-0.1/2})$$
$$= (11.5 - 1.0002 \cdot z_{0.95},\ 11.5 + 1.0002 \cdot z_{0.95})$$
$$= (11.5 - 1.0002 \cdot 1.64,\ 11.5 + 1.0002 \cdot 1.64)$$
$$= (9.86, 13.14)$$

Answers: hard disk failures

2. 95% CI means $\alpha = 0.05$

$$\begin{aligned}
&(11.5 - 1.0002 \cdot z_{1-0.05/2},\ 11.5 + 1.0002 \cdot z_{1-0.05/2}) \\
&= (11.5 - 1.0002 \cdot z_{0.975},\ 11.5 + 1.0002 \cdot z_{0.975}) \\
&= (11.5 - 1.0002 \cdot 1.96,\ 11.5 + 1.0002 \cdot 1.96) \\
&= (9.54, 13.46)
\end{aligned}$$

3. 99% CI means $\alpha = 0.01$

$$\begin{aligned}
&(11.5 - 1.0002 \cdot z_{1-0.01/2},\ 11.5 + 1.0002 \cdot z_{1-0.01/2}) \\
&= (11.5 - 1.0002 \cdot z_{0.995},\ 11.5 + 1.0002 \cdot z_{0.995}) \\
&= (11.5 - 1.0002 \cdot 2.33,\ 11.5 + 1.0002 \cdot 2.33) \\
&= (9.17, 13.83)
\end{aligned}$$

Answers: hard disk failures

- Notice: the confidence intervals get wider as the confidence level $1 - \alpha$ increases.
- None of these confidence intervals contains the manufacturer's target of 33.5 in. oz., so there is significant evidence that the process misses this target.
- Hence, there is a design flaw in the manufacturing process of the disk drives that must be corrected.

Controlling the width of a confidence interval

- If you want to estimate the breakaway torque with a 2-sided, 95% confidence interval with ± 2.0 in. oz. of precision, what sample size would you need?
- The confidence interval is:

$$
\begin{aligned}
&\left(\overline{x} - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}},\ \overline{x} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right) \\
&= \left(11.5 - z_{1-0.05/2}\cdot\frac{5.1}{\sqrt{n}},\ 11.5 + z_{1-0.05/2}\cdot\frac{5.1}{\sqrt{n}}\right) \\
&= \left(11.5 - z_{0.975}\cdot\frac{5.1}{\sqrt{n}},\ 11.5 + z_{0.975}\cdot\frac{5.1}{\sqrt{n}}\right) \\
&= (11.5 - 1.96\cdot 5.1\cdot n^{-1/2}, 11.5 + 1.96\cdot 5.1\cdot n^{-1/2}) \\
&= (11.5 - 9.996\cdot n^{-1/2}, 11.5 + 9.996\cdot n^{-1/2})
\end{aligned}
$$

Controlling the width of a confidence interval

The interval precision (half-width) δ is:

$$\begin{aligned}\delta &= \frac{1}{2}\left((11.5 + 9.996 \cdot n^{-1/2}) - (11.5 - 9.996 \cdot n^{-1/2})\right) \\ &= 9.996 \cdot n^{-1/2}\end{aligned}$$

We require δ to be at most 2:

$$\begin{aligned}2.0 &\leq 9.996 \cdot n^{-1/2} \\ n &\geq 25\end{aligned}$$

- We would need a sample of 25 disk drives to meet a precision of ± 2.0.