Introduction to formal statistical inference 6

Formal statistical inference uses probability theory to quantify the reliability of data-based conclusions. We want information on a population. We can use:

2. Interval estimates:

1

.

6.1 Large-sample confidence intervals for a mean

Many important engineering applications of statistics fit the following mold. Values for parameters of a data-generating process are unknown. Based on data, the goal is

Definition 6.1. A *confidence interval* for a parameter (or function of one or more parameters) is a data-based interval of numbers thought likely to contain the parameter (or function of one or more parameters) possessing a stated probability-based confidence or reliability.

A confidence interval is a realization of a **random interval**, an interval on the real line with a random variable at one or both of the endpoints.

Example 6.1 (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?

$$P(-1 in (2-2, 2+2)) = P(2-2 < -1 < 2+2)$$

$$= P(2-1 < 0 < 2+3)$$

$$\int = P(-1 < -2 < 3) = P(1 > 2 > -3)$$

$$= P(-3 < 2 < 1)$$

$$= P(2 \le 1) - P(2 \le -3)$$

$$= Q(1) - Q(-3)$$

$$= 0.84$$

Example 6.2 (More practice). Calculate:

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

 $P(2 \in (X - 1, X + 1)) = P(X - 1 < 2 < X + 1)$
 $= P(-1 < 2 - X < 1)$
 $= P(-1 < X - 2 < 1)$
 $= P(-1 < X - 2 < 1)$
 $= P(-1 < X - 2 < 1)$
 $= P(-0.5 < 2 < 0.5)$
 $= P(-0.5 < 2 < 0.5)$
 $= 0.6915 - 0.3085$
 $= 0.383$

•

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

$$P(6.1 \in (X-2, X+1)) = P(X-2 < 1.6 < X+1)$$

$$= P(-2 < 6.6 - X < 1)$$

$$= P(-1 < X-6.6 < 2)$$

$$= P(-1.4 \times -7 < 1.6)$$

$$= P(-1.4 \times -7 < 1.6)$$

$$= P(-\frac{1.4}{52} < \frac{X-7}{52} < \frac{1.6}{52})$$

$$= P(-0.9899 < 2 < 1.1313) - \frac{7}{2} \sim N(0,1)$$

$$= \Phi(1.13) - \Phi(-.98)$$

$$= 0.83708 - 0.1435 = 0.7073$$

Example 6.3 (Abstract random intervals). Let's say X_1, X_2, \ldots, X_n are iid with $n \ge 25$, mean μ , variance σ^2 . We can find a random interval that provides a lower bound for μ with $1 - \alpha$ probability:

Want a s.t.
$$P(\mu \in (a, \infty)) = 1 - \alpha$$

We know $X \sim N(M, \frac{C^2}{\sqrt{n}})$ by CLT

$$\Rightarrow (\overline{X} - \mu) \rightarrow N(0, 1) \quad by \quad standardization$$

$$Let = \frac{1}{2! - \infty} denote \quad he \quad 1 - \alpha \quad grantile \quad of \quad N(0, 1)$$

$$Zre N(0, 1), \quad P(Z = Z_{1-\alpha}) = 1 - \alpha \quad \text{for any } I - \alpha$$

$$Zre N(0, 1), \quad P(Z = Z_{1-\alpha}) = 1 - \alpha \quad \text{for any } I - \alpha$$

$$= \sum P\left(\frac{\overline{X}-\mu}{c^{2}/5n} \leq \overline{z}_{1-x}\right) \approx 1-x \\ P\left(\overline{X}-\mu \leq \overline{z}_{1-x}\frac{c^{2}}{5n}\right) \approx 1-x \\ P\left(\overline{X}-\overline{z}_{1-x}\frac{c^{2}}{5n} \leq \mu\right) \approx 1-x \\ P\left(\mu \in \left(\overline{X}-\overline{z}_{1-x}\frac{c^{2}}{5n} \leq \mu\right)\right) \approx 1-x \\ P\left(\mu \in \left(\overline{X}-\overline{z}_{1-x}\frac{c^{2}}{5n} \leq \mu\right)\right) \approx 1-x \\ \prod_{n=1}^{\infty} \prod_{i=1}^{\infty} \prod_{j=1}^{\infty} \prod_{i=1}^{\infty} \prod_{j=1}^{\infty} \prod_{j=1}^{$$

Calculate:

1. $P(\mu \in (-\infty, \overline{X} + z_{1-\alpha} \frac{\sigma}{\sqrt{n}})), X \sim N(\mu, \sigma^2)$ $= P(M < \tilde{X} + z_{1} + z_{2})$ $= P(-Z_{1-\alpha} - \frac{C}{2} < \overline{X} - m)$ $= \rho \left(- z_{1-x} < \frac{\overline{\chi} - M}{\kappa / r_{-}} \right)$ $\approx e(-z_{1x} < Z)$ Z~ N(0,1) by CLT if n325 = 1-a 1-d -0 2. $P(\mu \in (\overline{X} - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}, \overline{X} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}})), X \sim N(\mu, \sigma^2)$ $= P\left(\overline{X} - \overline{z_{1-u/2}} \sqrt{n} < M < \overline{X} + \overline{z_{1-u/2}} \sqrt{n}\right)$ $= p(-z_{1-\alpha/2} + \langle x_n - \chi \rangle \langle z_{1-\alpha/2} + \langle x_n \rangle)$ = p (-Z, v), 5 < X-M < Z, v) $= p\left(-\frac{1}{2}, -\frac{1}{2}, < \frac{\overline{X} - M}{6/\sqrt{2}} < \mathbf{Z}_{1} - \frac{1}{2}\right)$ $\approx P(-z_{1-kh} < Z < z_{1-kh})$ Z~N(0,1) by CLT for 7225 $= \Phi(z_{1-\alpha/2}) - \Phi(-z_{1-\alpha/2})$ 5= 1-a 0

AL

6.1.1 A Large-*n* confidence interval for μ involving σ

A $1 - \alpha$ confidence interval for an unknown parameter is the realization of a random interval that contains that parameter with probability $1 - \alpha$.

For random variables $X_1, X_2, \ldots, X_n^{\vee}$ iid with $E(X_1) = \mu$, $Var(X_1) = \sigma^2$, a $1 - \alpha$ confidence interval for μ is

n225

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

which is a realization from the random interval (has random random

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

• Two-sided $1 - \alpha$ confidence interval for μ

$$(\hat{\chi}_{-z_{1-w/2}} \frac{\epsilon}{\sqrt{n}}, \hat{\chi}_{+} z_{1-w/2} \frac{\epsilon}{\sqrt{n}})$$

 $\overline{\chi} \pm z_{1-w/2} \frac{\epsilon}{\sqrt{n}}$

• One-sided $1 - \alpha$ confidence interval for μ with a upper confidence bound

$$\left(-\infty\right)\overline{x} + z_{Ha}\frac{c}{\sqrt{n}}$$

• One-sided $1 - \alpha$ confidence interval for μ with a lower confidence bound

$$\left(\overline{\mathbf{x}}-\mathbf{z}_{1}, \mathbf{z}_{1}, \mathbf{z}_{1}\right)$$

•

Example 6.4 (Fill weight of jars). Suppose a manufacturer fills jars of food using a stable filling process with a known standard deviation of $\sigma = 1.6$ g. We take a sample of n = 47 jars and measure the sample mean weight $\overline{x} = 138.2$ g. A <u>two-sided</u> 90% confidence interval ($\alpha = 0.1$) for the true mean weight μ is:

$$(\overline{x} - z_{1-0!/2} \frac{c}{5\pi}) \overline{x} + z_{1-0!/2} \frac{c}{5\pi})$$

$$= (138.2 - z_{.95} \frac{1.6}{\sqrt{47}}) 138.2 + z_{.95} \frac{1.6}{\sqrt{47}})$$

$$= (138.2 - 1.64 (0.23), 138.2 + 1.64 (0.23))$$

$$= (137.82, 138.58)$$
Could also with as (38.2 ± 0.38)

Interpretation: ~We are 90% confident that he true wear fill is betreen 137.82 and 138.58g.

If we took 100 more samples of 47 jars each, roughly 90 of pose samples would produce confidence intervals containing the true men fill weight.

What if we just want to be sure that the true mean fill weight is high enough?

We could use a one-sided 90% CI with alover bound.

$$(\bar{x} - z_{1-\alpha} \frac{c}{\sqrt{n}}, \infty)$$

= (138.2 - $z_{,9} \frac{1.6}{\sqrt{47}}, \infty)$
= (138.2 - 1.28(0.23), ∞)
= (137.91, ∞)

We are 90% confident that the true mean fill weight is above 137.91. **Example 6.5** (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 link 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-side 90% onfidence interval for the true mean breakaway torque of the relevant type of Winchester drive.

 $\sigma = 5.1, \ \bar{X} = 11.5, \ n = 26, \ 1 - \alpha = 0.9 \Rightarrow \alpha = 0.1$ = $(\bar{X} - \bar{Z}_{1-\alpha/a} + \sqrt{n}, \ \bar{X} + \bar{Z}_{1-\alpha/a} + \sqrt{n})$ "margin of error" = $(11.5 - \bar{Z}_{0.95} + \sqrt{3n}, \ 1.5 + \bar{Z}_{0.95} + \sqrt{3n})$ = $(11.5 - 1.65(1.0602), \ 11.5 + 16.5(1.000a))$ = (9.85, 13.15)We are 90% confident that \mathcal{M} (true mean breakforming there use of Winchester drives) is between 9.85 and 13.15 in.

2. An analogous two-sided 95% confidence interval.

1-0

$$\begin{aligned} 1 - \alpha &= 0.95 \implies \alpha = 0.05 \\ &= (\tilde{X} - \tilde{Z}_{1-\alpha/2} \sqrt{n}, \tilde{X} + \tilde{Z}_{1-\alpha/2} \sqrt{n}) \\ &= (11.5 - \tilde{Z}_{0.975} \sqrt{\frac{5.1}{326}}, 11.5 + \tilde{Z}_{0.975} \sqrt{\frac{5.1}{326}}) \\ &= (9.54, 13.46) \\ &\text{We ore 957 confident that M (true mean break away torque of whichester drives) is blue 9.54 and 134.6 in.02.} \\ &\text{constant, } T \subset L \rightarrow wider \subset L \\ &= 0.94 \end{aligned}$$

CL= 95%, smar M.E (±2)

Example 6.6 (Width of a CI). 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?



we should need a sample of at least 25 disks to most have at teast a precision of 2 in 62.

6.1.2 A generally applicable large-n confidence interval for μ

Although the equations for a $1 - \alpha$ confidence interval is mathematically correct, it is severely limited in its usefulness because

if requires us to know $\underline{\sigma}$. It is unusual to have to estimate M (we a C.I.), but know σ in real life. If $n \ge 25$ and σ is unknown, $Z = \frac{\overline{X} - \mu}{S/\sqrt{n}}$, where instead of $s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2}$.

is still **approximately standard normally distributed**. So, you can replace σ in the confidence interval formula with the sample standard deviation, *s*.

•

• Two-sided $1 - \alpha$ confidence interval for μ

- One-sided 1α confidence interval for μ with a upper confidence bound (- ∞ , $\overline{X} + \overline{Z}_{1-\alpha}$, $\overline{S}_{1-\alpha}$)
- One-sided 1α confidence interval for μ with a lower confidence bound

$$(\bar{X} - \bar{Z}_{1-\alpha} \frac{s}{5n}, \infty)$$

Example 6.7. Suppose you are a manufacturer of construction equipment. You make 0.0125 inch wire rope and need to determine how much weight it can hold before breaking so that you can label it clearly. Here are breaking strengths, in kg, for 41 sample wires:

[1] 100.37 96.31 72.57 88.02 105.89 107.80 75.84 92.73 67.47 94.87 95.24 119.75 114.83 101.79 80.90 [11] 122.04 115.12 96.10 118.51 109.66 86.50 57.62 74.70 92.53 86.25 [21] 88.07 56.29 82.56 97.96 94.92 [31] 62.00 93.00 98.44 119.37 103.70 72.40 71.29 107.24 64.82 93.51 [41] 86.97

The sample mean breaking strength is 91.85 kg and the sample standard deviation is 17.6 kg. Using the appropriate 95% confidence interval, try to determine whether the breaking strengths meet the requirement of at least 85 kg.

 $\begin{aligned} 1-\alpha = 0.95 \Rightarrow a = 0.05 \\ \bar{x} = 91.85 \\ 5 = 17.6 \\ n = 41 \\ (\bar{x} - \bar{z}_{1-\alpha}, \frac{5}{\sqrt{n}}, \infty) \\ = (91.85 - \bar{z}_{0.95}, \frac{17.6}{\sqrt{41}}, \infty) \\ = (91.85 - 1.65, \frac{17.6}{\sqrt{41}}, \infty) \\ = (87.314, \infty) \\ \\ \text{ with } 95\% \text{ confidence, we have shown that } M \\ (\text{the true breaking strength of the wire ropes) is above $87.314 kg. \\ \end{aligned}$

> hence, the \$5 kg requirement is met w1 95%. confidence.

6.2 Small-sample confidence intervals for a mean

The most important practical limitation on the use of the methods of the previous sections is

That restriction comes from the fact that without it,

There is no way to conclude
$$\frac{X-M}{S/Sn} \sim N(0,1)$$

So, if one mechanically uses the large-*n* interval formula $\overline{x} \pm z \frac{s}{\sqrt{n}}$ with a small sample, there is no way of assessing what actual level if infidence (1-d) should be declared.

If it is sensible to model the observations as iid normal random variables, then we can arrive at inference methods for small-n sample means.

If his is true, $\frac{X-M}{S/\sqrt{n}}$ is not standard normal, But it is a named distribution.

6.2.1 The Student t distribution

Definition 6.2. The *(Student)* t distribution with degrees of freedom parameter ν is a continuous probability distribution with probability density

$$f(t) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{\pi\nu}} \left(1 + \frac{t^2}{\nu}\right)^{-(\nu+1)/2}$$
 for all t .

The t distribution

- is bell-shaped and symmetric about 0
- has fatter tails than the normal, but approaches the shape of the normal as $\nu \to \infty$.

We use the t table (Table B.4 in Vardeman and Jobe) to calculate quantiles.



traism we ares (n=6)

Example 6.8 (t quantiles). Say $T \sim t_5$. Find c such that $P(T \leq c) = 0.9$.

Table B.4

ν	Q(.9)	Q(.95)	Q(.975)	Q(.99)	Q(.995)	Q(.999)	Q(.9995)
1	3.078	6.314	12.706	31.821	63.657	318.317	636.607
2	1.886	2.920	4.303	6.965	9.925	22.327	31.598
3	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869

t Distribution Quantiles

Figure 1: Student's t distribution quantiles.

 $P(T \le 1.476) = 0.9$ O(p) for a ty is often denoted as typ. So, ts, o.9 = 1.476

6.2.2 Small-sample confidence intervals, σ unknown

If we can assume that X_1, \ldots, X_n are iid with mean μ and variance σ^2 , and are also normally distributed,

We cent use CLT since
$$n < 25$$
.
But we know $\frac{\overline{X} - M}{S/\sqrt{n}} \sim t_{n-1}$ (since $X_{1,...,X_{n}}$, iid $N(4,\sigma^{2})$)
Side note: If we do know σ , then $\frac{\overline{X} - M}{\sigma/\sqrt{n}} \sim N(0,1)$
even for small n (if $X_{1,...,X_{n}}$ iid $N(M, \sigma^{2})$)

We can then use $t_{n-1,1-\alpha/2}^{\alpha}$ instead of $z_{1-\alpha/2}$ in the confidence intervals.

• Two-sided $1 - \alpha$ confidence interval for μ

• One-sided $1 - \alpha$ confidence interval for μ with a upper confidence bound

 $(-\infty, \overline{X} + t_{n-1, 1-\alpha}, \frac{s}{\sqrt{n}})$

- One-sided $1 - \alpha$ confidence interval for μ with a lower confidence bound

$$(\bar{X} - t_{n-1,1-\alpha}, s_n, \infty)$$

Example 6.9 (Concrete beams). 10 concrete beams were each measured for flexural strength (MPa). Assuming the flexural strengths are iid normal, calculate and interpret a two-sided 99% CI for the flexural strength of the beams.

[1] 8.2 8.7 7.8 9.7 7.4 7.8 7.7 11.6 11.3 11.8

$$n = 10, \alpha = 0.01$$

 $x = \frac{1}{10} (s.2 + ... + 11.9) = 9.2$
 $S = \sqrt{\frac{1}{9} [(8.2 - 9.2)^2 + ... + (11.8 - 9.2)^2]} = 1.76$
Two-sided 99./ CI: $(x - t_{n-1,1} - \alpha/2 \frac{s}{5n}, x + t_{n-1,1} - \alpha/2 \frac{s}{5n})$
 $= (9.2 - t_{9,0.915} \frac{1.76}{15}, 9.2 + t_{9,0.915} \frac{1.76}{150})$
 $= (9.2 - 3.250 (0.55c), 9.2 + 3.250 (0.55c))$
 $(7.393, 11.007)$
We are 99./ confident that M (true mean flexural
strength of this kind of concrete beam) is thus
7.393 Mpa and 11.007 Mpa.

Is the true mean flexural strength below the minimum requirement of 11 MPa? Find out

with the appropriate 95% CI.

 $= (-\infty, \overline{X} + t_{n-1, 1-\alpha} \frac{\sqrt{3}}{\sqrt{16}})$ = (-\infty, 9.2 + t_{1, 0.95} \frac{1.76}{\sqrt{16}}) = (-\infty, 9.2 + 1.833 \frac{1.76}{\sqrt{16}}) = (-\infty, 10.21) what if we accidentally tried to solve this problem using a lower-bound cI instead of an upper-bound CI? (X-En-1,1-d Tn, ∞) This would be a problem b/c our answer would be (c, ∞), where c is answer would be (c, ∞), where c is a constant lower-bound. A part of integral (c, ∞) will always contain estimates of M (c, ∞) will always contain estimates of M preater than II (since ∞) II), and the preater than II (since ∞) II), and the minimal requirement of II will always pass.

we are 95% confident that M (true mean Flexible) strength of this kind of concrete beam) is below 10.21 MPQ. (That's below 11 MPQ) At d = 0.05, we have shown the true mean Flexible strength is < 11 MPQ, and the minimal requirement is not met. **Example 6.10** (Paint thickness). Consider the following sample of observations on coating thickness for low-viscosity paint.

[1] 0.83 0.88 0.88 1.04 1.09 1.12 1.29 1.31 1.48 1.49 1.59 1.62 1.65 1.71 [15] 1.76 1.83



A normal QQ plot shows that they are close enough to normally distributed.

Calculate and interpret a two-sided 90% confidence interval for the true mean thickness. n=16, q=0.1 $\bar{x} = \frac{1}{16}(0.83 + \dots + 1.83) = 1.35$ mm $5 = \sqrt{15}((0.83 - 1.35)^2 + \dots + (1.83 - 1.35)^2 = 0.34$ mm $= (\bar{x} - t_{n-1,1-\alpha/2}, \bar{x}, \bar{x} + t_{n-1,1-\alpha/2}, \bar{x}, \bar{x})$ $= (1.35 - t_{15}, 0.15, \sqrt{16}, 1.35 + t_{15}, 0.15, \sqrt{16})$ $= (1.35 - 1.753, \frac{0.34}{\sqrt{16}}, 1.35 + t_{15}, 0.15, \sqrt{16})$ = (1.201, 1.499)

WE are 90% confident that M (true mean thickness) is Wan 1.201 and 1.499 mm.

	ibution Qu	untiles					
ν	Q(.9)	Q(.95)	Q(.975)	Q(.99)	Q(.995)	Q(.999)	Q(.9995)
1	3.078	6.314	12.706	31.821	63.657	318.317	636.607
2	1.886	2.920	4.303	6.965	9.925	22.327	31.598
3	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	1.345	1.701	2.145	2.624	2.977	3.787	4.140
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073
			1				
16	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	1.325	1.725	2.086	2.528	2.845	3.552	3.849
21	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	1.296	1.671	2.000	2.390	2.660	3.232	3.460
120	1.289	1.658	1.980	2.358	2.617	3.160	3.373
∞	1.282	1.645	1.960	2.326	2.576	3.090	3.291

Table B.4t Distribution Quantiles

This table was generated using MINITAB.