1-SAMPLE INFERENCE: INTRO TO CONFIDENCE INTERVALS [DEVORE 7.1]

- **INTERVAL ESTIMATORS & CONFIDENCE INTERVALS (DEFINITION):**

Given a population with parameter θ and random sample $\mathbf{X} := (X_1, \dots, X_n)$.

A $100(1-\alpha)\%$ **interval estimator** $(\theta_{L,1-\alpha}(\mathbf{X}),\ \theta_{U,1-\alpha}(\mathbf{X}))$ for θ is an interval of values constructed from the random sample such that:

$$\mathbb{P}\left[\theta_{L,1-\alpha}(\mathbf{X}) < \theta < \theta_{U,1-\alpha}(\mathbf{X})\right] = 1-\alpha$$

Moreover, suppose a sample $\mathbf{x} := (x_1, \dots, x_n)$ is taken from the population.

The $100(1-\alpha)\%$ **confidence interval** $(\theta_{L,1-\alpha}(\mathbf{x}),\ \theta_{U,1-\alpha}(\mathbf{x}))$ for θ is the $100(1-\alpha)\%$ interval estimator but replacing each parameter involved in $\theta_{L,1-\alpha}(\mathbf{X})$ & $\theta_{U,1-\alpha}(\mathbf{X})$ with its point estimator.

The percent $100(1-\alpha)\%$ is called the **confidence level**.

- **WRONG INTERPRETATIONS OF A CONFIDENCE INTERVAL:**

"The probability that μ is between 9.42 years & 15.77 years is 0.95"

"There is a 95% **chance** that μ is between 9.42 years & 15.77 years."

WHY ARE THESE INTERPRETATIONS WRONG????

Because the population parameter μ is **not random**!!!

i.e. μ is not changing, μ is some fixed value – we just don't know that value!

Now, the sample mean $\overline{X}$ of a random sample from this population is random.

Then, it can be shown that the interval estimator involved the following probability:

$$\mathbb{P}(\overline{X} - 3.325 < \mu < \overline{X} + 3.325) = 0.95$$

So, in effect, what is actually random is the interval itself!!!!

- **RIGHT INTERPRETATIONS OF A CONFIDENCE INTERVAL:**

"There is 95% **confidence** that μ is between 9.42 years and 15.77 years."

"The probability that the next sample's computed CI will contain μ is 0.95"

"After taking many many samples from the population, about 95% of the resulting CI's will contain μ."

- **CONSTRUCTING CONFIDENCE INTERVALS (PROCEDURE):**

GIVEN: Random sample $\mathbf{X} := (X_1, \dots, X_n)$ of a population with parameter θ.

TASK: Construct the $100(1-\alpha)\%$ Confidence Interval for θ. $(0 < \alpha < 1)$

(1) Produce a **pivot** $Q(\mathbf{X};\theta)$ which is a statistic such that:

– Q is a function of both random sample $X_1, \dots, X_n$ and parameter θ.

– The pdf of Q, $f_Q(x)$, does not depend on θ or any other parameters.

(2) Find real numbers $a < b$ such that the following probability holds:

$$\mathbb{P}[a < Q(\mathbf{X};\theta) < b] = 1-\alpha$$

(3) Manipulate the above inequalities to isolate parameter θ:

$$\mathbb{P}\left[\theta_{L,1-\alpha}(\mathbf{X}) < \theta < \theta_{U,1-\alpha}(\mathbf{X})\right] = 1-\alpha$$

(4) The $100(1-\alpha)\%$ interval estimator for θ is: $\theta \in (\theta_{L,1-\alpha}(\mathbf{X}),\ \theta_{U,1-\alpha}(\mathbf{X}))$

(5) Obtain a sample $\mathbf{x} := (x_1, \dots, x_n)$ from the population

(6) Compute point estimator(s) for each parameter in $\theta_{L,1-\alpha}(\mathbf{X})$ & $\theta_{U,1-\alpha}(\mathbf{X})$

(7) Replace the each parameter with its point estimate, resulting in the CI

$$(\theta_{L,1-\alpha}(\mathbf{x}),\ \theta_{U,1-\alpha}(\mathbf{x}))$$

EX 7.1.1: Consider the Normal(μ, σ^2) population of all college student heights and the average height μ (in feet.)

(a) Construct a 95% interval estimator for μ.

(b) Construct a 90% interval estimator for μ.

(c) Suppose a sample of size ($n = 10$) is taken from the population.
Moreover, the sample size $\overline{x} = 5.1$ ft and the std deviation $s = 1.2$ ft.
Construct the 90% and 95% confidence intervals for μ.

(d) Suppose a sample of size ($n = 20$) is taken from the population.
Moreover, the sample size $\overline{x} = 5.5$ ft and the std deviation $s = 1.4$ ft.
Construct the 90% and 95% confidence intervals for μ.

(e) Which of the four confidence intervals from parts (c) & (d) has the most reliability and the most precision?