Lecture 25

Possible Outcomes

	H_0 True	H_1 True.
Accept H_0	✓	Type II error
Reject H_0	Type I error	✓

Define: $\alpha = P(\text{Type I error}) :=$ "significance level" of the test.

$\beta = P(\text{Type II error})$, $1-\beta$ is the "power" of the test.

$1-\beta =$ prob. of correctly rejecting H_0.

Example: Suppose we sample a normal distribution with $\sigma = 10$, with sample size 35.

Suppose we have
$$H_0: \mu = 50$$
$$H_1: \mu \neq 50$$

Suppose our critical region is $\bar{x} \geq 52$, $\bar{x} \leq 48$.

a) Find α.

b) Suppose $\mu = 53$. What is β?

a) $\alpha = P(\text{Reject } H_0,\ H_0 \text{ true}) = P(\bar{X} \geq 52 \text{ or } \bar{X} \leq 48 \text{ and } \mu = 50)$

$$= P(\bar{X} \geq 52,\ \mu = 50) + P(\bar{X} \leq 48,\ \mu = 50)$$

$$= \underbrace{P\left(Z \geq \frac{52-50}{10/\sqrt{35}}\right) + P\left(Z \leq \frac{48-50}{10/\sqrt{35}}\right)}_{\text{z scores.}} = 0.238.$$

b) $\beta = P(\text{accept } H_0 \text{ given that } H_0 \text{ is false})$

if $\mu = 53$, then

$$\beta = P(\text{accept } H_0 \text{ given that } \mu = 53)$$

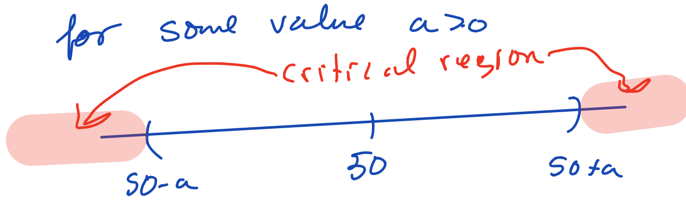
$$= P\left(\frac{48-53}{10/\sqrt{35}} \leq Z \leq \frac{52-53}{10/\sqrt{35}}\right) = 0.274.$$

- The concept of "significance" gives a good way to quantitatively determine a critical region.

- This gives a convenient way to report the outcome of a hypothesis test is to state whether H_0 was accepted or rejected at a certain level of significance ("fixed significance level testing")

<u>Ex:</u> Suppose we again sample $N(\mu, (10)^2)$ 35 times and that $H_0: \mu = 50$, $H_1: \mu \neq 50$. H_1 is two-sided.
Find a (symmetric) critical region if $\alpha = 0.05$.

<u>Solution:</u> We suppose that our critical values are $50 \pm a$ for some value $a > 0$



- As in the last example, we have

$$0.05 = \alpha = P(\text{Reject } H_0, H_0 \text{ true}) = P([\bar{X} \le 50-a] \text{ or } [\bar{X} \ge 50+a]) \text{ and } \mu = 50)$$

$$= P\left(N(50, \tfrac{10^2}{35}) \le 50-a\right) + P\left(N(50, \tfrac{10^2}{35}) \ge 50+a\right)$$

$$= P\left(Z \le \frac{-a}{10/\sqrt{35}}\right) + P\left(Z \ge \frac{a}{10/\sqrt{35}}\right)$$

$$= 2P\left(Z \le \frac{-a}{10/\sqrt{35}}\right)$$

So $$0.025 = P\left(Z \le \frac{-a}{10/\sqrt{35}}\right) = \Phi\left(\frac{-a}{10/\sqrt{35}}\right)$$

look up Z-scores ...... $\Phi(-1.96) = 0.025$

So $$-1.96 = \frac{-a}{10/\sqrt{35}} \Rightarrow a = (1.96)(10/\sqrt{35}) \approx 3.313.$$

- This gives a critical region of $\bar{x} \le 46.687$, $\bar{x} \ge 53.313$.

- Note that essentially the same process can be used to determine a 1-sided critical region for the test (eg $\alpha = P(\bar{X} \le 50+a)$)

P-values

- There is something a little unsatisfactory about constructing the critical region with a given, fixed significance α.
- In the last example, our H_0 was $\mu = 50$; given a sample size of 35, we found a region such that we accept H_0 if $\bar{x} \in (46.687, 53.313)$.
- Suppose $\bar{x} = 53.3129999$. Then we accept H_0. But is that value that much more likely than 53.313

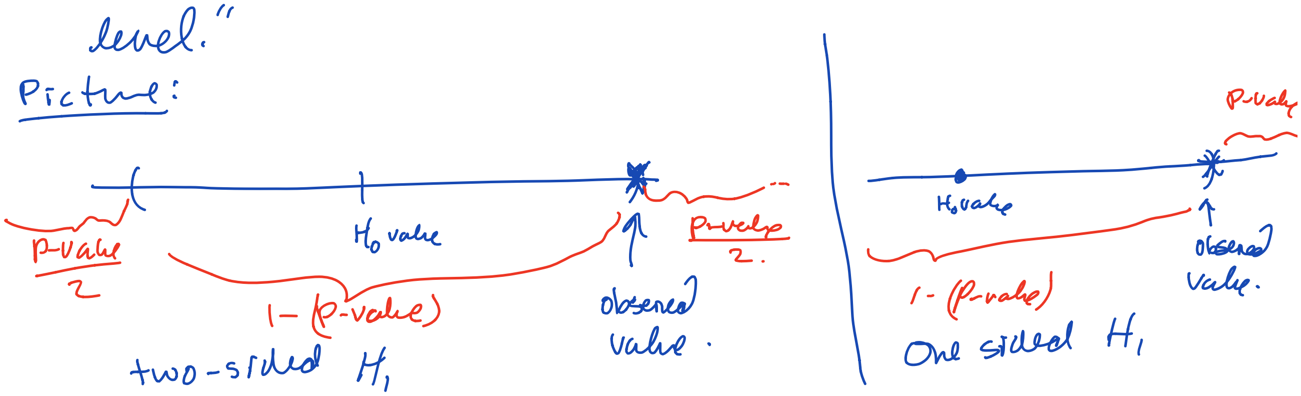
(for which we would reject H_0)?

- Rather than just fixing a significance level α, <u>P-values</u> give a more dynamic approach to hypothesis testing.

✱ Given an experiment with null hypothesis H_0, the <u>P-value</u> of the experiment is the smallest significance level that would lead to a rejection of H_0 with the given data. ✱

- the P-value is sometimes called the "observed significance level."

<u>Picture:</u>



<u>Ex:</u> Again, we sample a normal distribution $N(\mu, 10^2)$, $n = 35$ times. Suppose $H_0: \mu = 50$. If our observation is $\bar{x} = 54$ ← observed data.

a) Find the p-value if $H_1: \mu \neq 50$.

Observe that the extreme values around 50 are $50 \pm |54 - 50|$, so 50 ± 4.

So the p-value is $2 \cdot P(\bar{X} \leq \underbrace{46}_{50-4} \mid \mu = 50)$

$$P(\bar{X} \leq 46 \mid \mu = 50) = P\left(\underbrace{\frac{\bar{X} - 50}{10/\sqrt{35}}}_{Z} \leq \frac{46 - 50}{10/\sqrt{35}}\right)$$

$$= 0.009$$

So the P-value is $2(0.009) = 0.018$

i.e. getting 54 or 46 is fairly unlikely, even if H_0 is true.

b) Find the P-value if $H_1: \mu > 50$.

Then p-value $= P(\overline{X} \geq 54 \text{ given } \mu = 50)$

$$= P\left(\frac{\overline{X} - 50}{10/\sqrt{35}} \geq \frac{54-50}{10/\sqrt{35}}\right) = 1 - \Phi\left(\frac{4}{10/\sqrt{35}}\right)$$

$$= 0.009.$$

Motto: The p-value is a measure of the risk that we make an incorrect decision if we reject Ho.

- So for example, if we are given a value of our test statistic and it has a very high p-value, then we are at high risk of a Type I error if we reject Ho.

- Notice that p-values give us a refinement of the "fixed significance level" testing.
  - ↳ - fix a significance level α.
    - if the observed value is in the critical region, then the p-value is smaller than α
      - ↳ if p-value $\leq \alpha$, Reject Ho.
    - otherwise, accept H_0, report the result with p-value.

Relationship between Hypothesis Test and Confidence Intervals

- close relationship between a hypothesis test for a parameter θ and the confidence interval for θ.

- suppose that $[L, U]$ is a $100(1-\alpha)\%$ confidence interval for θ

- Then a hypothesis test of significance level α for the hypothesis

$$H_0: \theta = \theta_0 \ , \quad H_1: \theta \neq \theta_0$$

(θ_0 some number)

will lead to a rejection of H_0 if and only if $\theta_0 \notin [L, U]$.

- This gives an equivalent way of performing fixed sig. level testing:
  - Given a hypothesis $H_0: \theta = \theta_0 \quad H_1: \theta = \theta_0$, choose a test statistic $\hat{\Theta}$ for θ and a sig. level α.
  - for a given point estimate $\hat{\Theta} = \hat{\theta}$, construct the CI around $\hat{\theta}$.
    - If θ_0 is in the interval, accept H_0.
    - If θ_0 is not in the interval, reject H_0.