A general approach ~~for~~ to get a level α test

- Estimate θ by its point estimator $\hat{\theta}$
- Compute s.e.$(\hat{\theta})$ assuming $\theta = \theta_0$. Estimate it if it's unknown.
- Compute a **test statistic** T that measures measures how consistent the data are with H_0. Often, T has the form:

$$T = \frac{\hat{\theta} - \theta_0}{\widehat{se}(\hat{\theta})}.$$

- Find the **null distribution** — the distribution of T assuming H_0 is true.
- Find the form of the **rejection region** $\mathcal{R}$ — the set of values of T for which H_0 is rejected.
- **Acceptance region** $\mathcal{A}$ = Complement of $\mathcal{R}$.
- Determine $\mathcal{R}$ by ensuring that the level of significance of the test is α, i.e., $P(\text{reject } H_0 | H_0 \text{ is true}) = \alpha$.

Some common rejection regions

Suppose $T = \dfrac{\hat{\theta} - \theta_0}{\widehat{se}(\hat{\theta})}$.

In this case, it is often easy to guess $\mathcal{R}$.

Case 1: $H_0 : \theta = \theta_0$ against $H_1 : \theta \neq \theta_0$

Reject H_0 if $|T| > c$ ← positive cutoff.

Case 2: $H_0 : \theta = \theta_0$ against $H_1 : \theta > \theta_0$

Reject H_0 if $T > c$ ← positive cutoff.

Case 3: $H_0 : \theta = \theta_0$ against $H_1 : \theta < \theta_0$

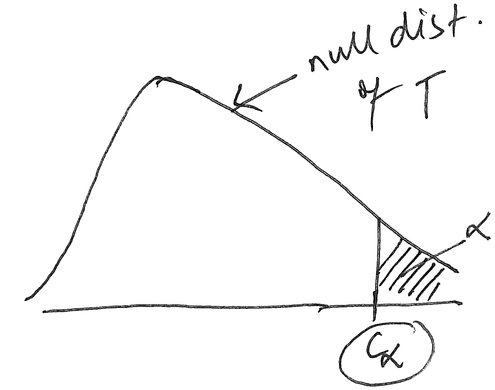
Strong evidence against H_0 [i.e., in favor of H_1] if T is too small, because in this case $\hat{\theta} \ll \theta_0$, Reject H_0 if $T < c$. ← an appropriate negative cutoff.

Compute the critical point in a way that ensures that the level of the test equals the prescribed α.

The corresponding level α tests:

Suppose c_α is such that $P(T > c_\alpha | \theta = \theta_0) = \alpha$.

H_0 is true.



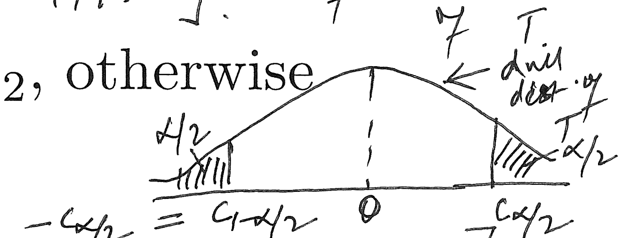
$(1-\alpha)$-th percentile of null dist.

Case 1: $H_0 : \theta = \theta_0$ against $H_1 : \theta \neq \theta_0$ [Recall: Reject when $|T| > c$].

$\mathcal{R} = \{|T| > c_{\alpha/2}\}$, i.e., reject H_0 when $|T| > c_{\alpha/2}$, otherwise accept it.

we need to get c by ensuring that

$P[\,|T| > c \,|\, H_0 \text{ is true}] = \alpha$.

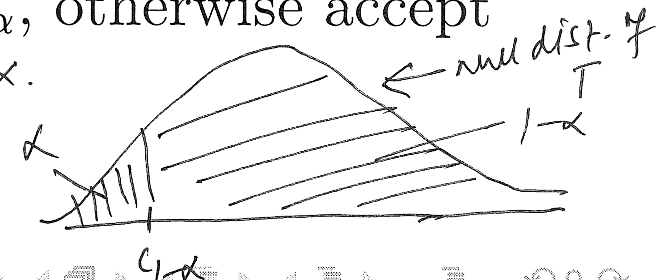


Case 2: $H_0 : \theta = \theta_0$ against $H_1 : \theta > \theta_0$ [Recall: Reject when $T > c$].

$\mathcal{R} = \{T > c_\alpha\}$, i.e., reject H_0 when $T > c_\alpha$, otherwise accept it.

Case 3: $H_0 : \theta = \theta_0$ against $H_1 : \theta < \theta_0$ [Recall: Reject when $T < c$] negative

$\mathcal{R} = \{T < \cancel{wc_\alpha}\ c_{1-\alpha}\}$, i.e., reject H_0 when $T < \cancel{wc_\alpha}\ c_{1-\alpha}$, otherwise accept it.



$c_{1-\alpha}$ is such that

$P[\,T < c_{1-\alpha} \,|\, H_0 \text{ is true}] = \alpha$.

[Note: If dist. of T is symmetric about zero, then $c_{1-\alpha} = -c_\alpha$].

Recap

- H_0 and H_1 — Null and alternative hypothesis
- Type I and type II errors
- Test statistic, Null distribution of test statistic
- Level α test: $P[\text{Reject } H_0 \mid H_0 \text{ is true}] \leq \alpha$

'=' in this course

$P[\text{Type I error}]$

T_{obs} = observed value of T.

Hypothesis testing (continued)

We can perform a level α test by comparing T_{obs} with the critical point. But how strong is the evidence against the null? This is formally measured by *p-value*. Let's play a game to motivate its definition.

My bag has 10 small balls. I claim that 8 are red and 2 are ~~blue~~ black. I will bet 3 people a candy bar that a ~~blue~~ black ball will come up. My chances are not very good but I will take them anyway.

trial #	PKC vs ?	color drawn	winner
1	Ali	black	PKC
2	Anton	black	PKC
3	Giri	black	PKC

Q. Does is it seem reasonable that I would win . . . 3 times in 3 trials if the bag contained 2 ~~blue~~ black balls?

$\Rightarrow$ data don't seem to be consistent with claim

$\Rightarrow$ claim is not true.

Plausible but chance of this very small.

Let's cast this problem as a test of hypothesis.

Hypotheses:

$H_0: p = 0.20$ ← prob of black balls in the bag

$H_1: p \neq 0.20$

T and T_{obs}

$T = X_1 + X_2 + X_3$ = # black balls drawn.

$T_{obs} = 3$.

Null distribution T:

$T \sim \text{Bin}(n=3, p=0.20)$.

Q. What is the actual chance of getting T_{obs} if H_0 is true? What does it indicate about H_0?

$$P[T = T_{obs} = 3 \mid H_0 \text{ is true}] = P[\text{Bin}(n=3, p=0.20) = 3]$$
$$= (0.2)(0.2)(0.2) = 0.008$$

⇒ Data that we got is 'rare' under H_0 ⇒ H_0 is not true. ⇒ Reject H_0.

p-value: The probability of getting a T that is *as extreme or more extreme* than T_{obs} assuming that H_0 is true.

- Small p-value implies T_{obs} is 'rare' if H_0 is true
  $\Rightarrow$ indicates that H_0 is not true, and we should reject H_0.
- Smaller the p-value, stronger the evidence against H_0.
- Level α test: Reject H_0 if p-value $\leq \alpha$.
- Another interpretation of p-value: The smallest level of significance at which H_0 is rejected.
- Advantage of p-value over critical point: Using p-value is more informative than simply using the critical pt.

Q. Is p-value $= P(H_0 \text{ is true})$?

A. NO.

- H_0 is either true or not true, but we don't know the truth. Certainly, H_0 is not a random quantity.
- p-value tells us how likely our T_{obs} is (or something more extreme) if H_0 is true.

Summary of steps in a hypothesis test:

- Formulate H_0 and H_1
- Find a test statistic T and get its null distribution
- Compute T_{obs}
- Use the null distribution to compute either the critical point or the p-value for the test.
- State your conclusion. (in lay man terms also).

  don't just say Ho is accepted or rejected.

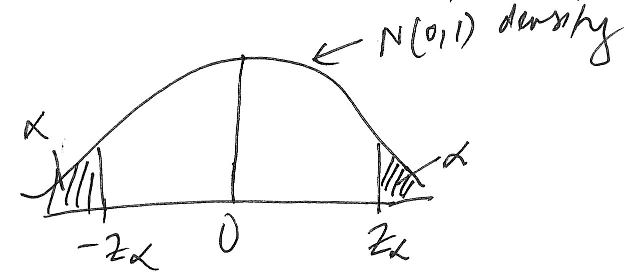
Some specific tests
One-sample tests for μ where $X \sim N(\mu, \sigma^2)$

NOTE: when H_0 is true: $\bar{X} \sim N\left[\mu = \mu_0, \frac{\sigma^2}{n}\right]$.

Case 1: z**-test (known** σ^2**):** $H_0 : \mu = \mu_0$

Test statistic: General: $T = \frac{\hat{\theta} - \theta_0}{se(\hat{\theta}_0)}$ Here: $z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$

$\sim N(0,1)$ when H_0 is true.



Critical point for the level α test:

One-sided alternative: z_α or $-z_\alpha$

right-tailed left-tailed

Two-sided alternative:

$z_{\alpha/2}$