General

- The Procedure of Hypotheses Testing, Please specify the five steps explicitly.

Step 1: State the hypotheses. Rules:

(a) H_0 : Status quo, or no relationship, or no difference. In most situations, the researcher hopes to disprove or reject the null hypothesis. "=" always goes with H_0

(b) H_1 : The assumed status quo is false, or that there is a relationship, or that there is a difference. In most situations, this hypothesis is what the researcher hopes to prove. Never use "=" or "≤", "≥". H_0 and H_1 should be contradicting.

Step 2: Select test statistics $T(X_1, \ldots, X_n)$

Step 3: Get its corresponding null distribution for T. That is the distribution of T under H_0

Step 4: Determine the rejection/critical region C (or calculate the p-value)

Step 5: Make a decision : If the sample value of T does fall in reject region C, we reject H_0; otherwise we fail to reject H_0. Statements:

- If reject H_0, "There is evidence / The data supports that (plug in H_1) at the significance level α"
- If fail to reject H_0, "There is not enough evidence to reject the null that (plug in H_0) at the significance level α "

- Type I error $= P($ Reject $H_0|H_0$ is true)
- Type II error $= P($Fail to reject $H_0|H_0$ is false)

Tests about population Mean

1. Case I: Normal population with known σ

   - Null hypothesis: $H_0 : \mu = \mu_0$
   - Test statistics value: $z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$

H_1	Rejection Region for level α	Type II error $\beta(\mu')$
$H_1 : \mu > \mu_0$	$z \geq z_\alpha$	$\Phi(z_\alpha + \frac{\mu_0-\mu'}{\sigma/\sqrt{n}})$
$H_1 : \mu < \mu_0$	$z \leq -z_\alpha$	1- $\Phi(-z_\alpha + \frac{\mu_0-\mu'}{\sigma/\sqrt{n}})$
$H_1 : \mu \neq \mu_0$	$z \geq z_{\alpha/2}$ or $z \leq -z_{\alpha/2}$	$\Phi(z_{\alpha/2} + \frac{\mu_0-\mu'}{\sigma/\sqrt{n}}) - \Phi(-z_{\alpha/2} + \frac{\mu_0-\mu'}{\sigma/\sqrt{n}})$

2. Case II: population with unknown σ and sample size $n > 40$

   - Null hypothesis: $H_0 : \mu = \mu_0$
   - Test statistics value: $z = \frac{\bar{x}-\mu_0}{s/\sqrt{n}}$

H_1	Rejection Region for level α	Type II error $\beta(\mu')$
$H_1 : \mu > \mu_0$	$z \geq z_\alpha$	$\Phi(z_\alpha + \frac{\mu_0-\mu'}{s/\sqrt{n}})$
$H_1 : \mu < \mu_0$	$z \leq -z_\alpha$	1- $\Phi(-z_\alpha + \frac{\mu_0-\mu'}{s/\sqrt{n}})$
$H_1 : \mu \neq \mu_0$	$z \geq z_{\alpha/2}$ or $z \leq -z_{\alpha/2}$	$\Phi(z_{\alpha/2} + \frac{\mu_0-\mu'}{s/\sqrt{n}}) - \Phi(-z_{\alpha/2} + \frac{\mu_0-\mu'}{s/\sqrt{n}})$

3. Case III: population with unknown σ and sample size $n \leq 40$

   - Null hypothesis: $H_0 : \mu = \mu_0$
   - Test statistics value: $t = \frac{\bar{x}-\mu_0}{s/\sqrt{n}}$

H_1	Rejection Region for level α
$H_1 : \mu > \mu_0$	$t \geq t_{\alpha,n-1}$
$H_1 : \mu < \mu_0$	$t \leq -t_{\alpha,n-1}$
$H_1 : \mu \neq \mu_0$	$t \geq t_{\alpha/2,,n-1}$ or $t \leq -t_{\alpha/2,n-1}$

Tests Concerning a Population Proportion

1. Large sample test for one sample proportion.

   - Conditions: $np_0 \geq 10$ AND $n(1-p_0) \geq 10$
   - Null Hypothesis: $H_0 : p = p_0$
   - Test statistic value: $z = \dfrac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}}$

H_1	Rejection Region for level α	Type II error $\beta(\mu')$
$H_1 : p > p_0$	$z \geq z_\alpha$	$\beta(p') = P\left(z < \dfrac{p_0 - p' + z_\alpha\sqrt{p_0(1-p_0)/n}}{\sqrt{p'(1-p')/n}}\right)$
$H_1 : p < p_0$	$z \leq -z_\alpha$	$\beta(p') = 1 - P\left(z < \dfrac{p_0 - p' - z_\alpha\sqrt{p_0(1-p_0)/n}}{\sqrt{p'(1-p')/n}}\right)$
$H_1 : p \neq p_0$	$z \geq z_{\alpha/2}$ or $z \leq -z_{\alpha/2}$	$\beta(p') = P\left(z < \dfrac{p_0 - p' + z_{\alpha/2}\sqrt{p_0(1-p_0)/n}}{\sqrt{p'(1-p')/n}}\right)$ $-P\left(z < \dfrac{p_0 - p' - z_{\alpha/2}\sqrt{p_0(1-p_0)/n}}{\sqrt{p'(1-p')/n}}\right)$

2. Small sample test for one sample proportion. (One sided only)

   Notation: $B(x, n, p) \doteq P(X \leq x | X \sim \text{Binomial}(n, p))$

   - $H_0 : p = p_0$
   - Test statistics X, the number of success events.
   - The corresponding null distribution for X: under H_0, $X \sim \text{Binomial}(n, p_0)$.

(a) $H_1 : p < p_0$:

- Rejection region: of the form $\{x : x \leq c\}$. The critical value c satisfies that $B(c, n, p_0) \leq \alpha$ and $B((c+1), n, p_0) > \alpha$.
- Type II error: If the true value of p is $p' < p_0$, then the type II error is

$$\begin{aligned} \beta(p') &= P[\text{Fail to reject } \ H_0 \text{ when } \ X \sim \text{Binomial}(n, p')] \\ &= P[X \geq (c+1) \ \text{ when } X \sim \text{Binomial}(n, p')] \\ &= 1 - B(c; n, p') \end{aligned}$$

(b) $H_1 : p > p_0$

- Rejection region: of the form $\{x : x \geq c\}$. The critical value c satisfies that $1 - B((c-1), n, p_0) \leq \alpha$ and $1 - B(c, n, p_0) > \alpha$.
- Type II error: If the true value of p is $p' > p_0$, then the type II error is

$$\begin{aligned} \beta(p') &= P[\text{Fail to reject } \ H_0 \text{ when } \ X \sim \text{Binomial}(n, p')] \\ &= P[X < c \ \text{ when } X \sim \text{Binomial}(n, p')] \\ &= B((c-1); n, p') \end{aligned}$$

P-values

- Definition: **P-value** is the probability, of obtaining a test statistic at least as contradictory to the null hypothesis as the one we have calculated from the available sample, assuming the null hypothesis is true.
- Decision based on p-values:
  - If p-value $\leq \alpha$ we reject H_0
  - If p-value $> \alpha$ we do not reject H_0
- p-value calculation: Suppose the test statistic is Z (or T), and the corresponding value in the sample is z (or t). Then the p-value is
  1. Two-sided test $H_1 : \mu \neq \mu_0$: $P(|Z| > z)$, $P(|T| > t)$;
  2. Upper-tailed test $H_1 : \mu > \mu_0$: $P(Z > z)$, $P(T > t)$;
  3. Lower-tailed test $H_1 : \mu < \mu_0$: $P(Z < z)$, $P(T < t)$.