Hypotheses Testing:

- A hypothesis is a statement about one or more populations.
- A statistical hypothesis is a conjecture (or a statement) concerning the population which can be evaluated by appropriate statistical technique.
- We usually test the null hypothesis (H_0) against the alternative (or the research) hypothesis (H_A or H_1) by choosing one of the following situations:

  Two-sided hypothesis:

  H_0: $\theta = \theta_0$ against H_A: $\theta \neq \theta_0$

  One-sided hypothesis:

  (i) H_0: $\theta = \theta_0$ against H_A: $\theta < \theta_0$

  (ii) H_0: $\theta = \theta_0$ against H_A: $\theta > \theta_0$
- There are 4 possible situations in testing a statistical hypothesis:

		Condition of Null Hypothesis H_0 (Nature/Reality)	
		H_0 is true	H_0 is false
Possible Action (Decision)	Accepting H_0	Correct Decision	Type II error (β)
	Rejecting H_0	Type I error (α)	Correct Decision

- There are two types of errors:

  (i) Type I error = Rejecting H_0 when H_0 is true

  P(Type I error) = P(Rejecting H_0 | H_0 is true) = α

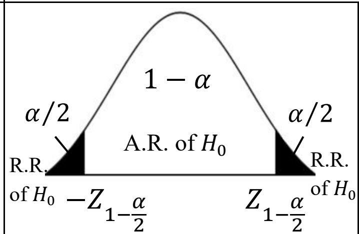
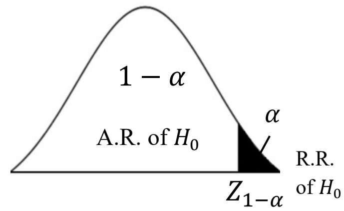
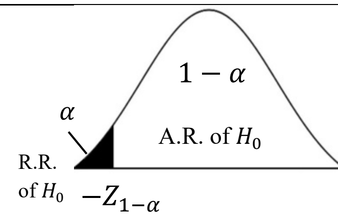
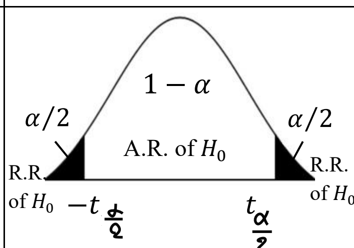
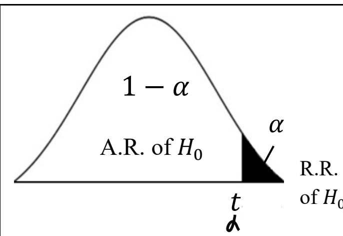
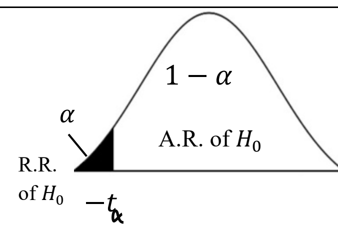
  Which is called the significance level of the test.

  (ii) Type II error = Accepting H_0 when H_0 is false

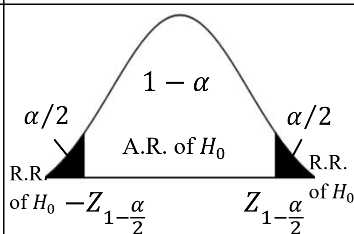
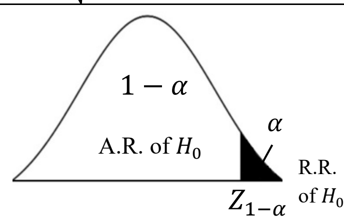
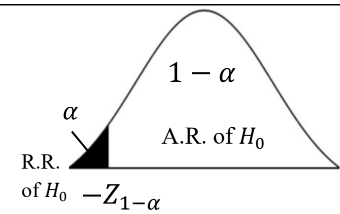
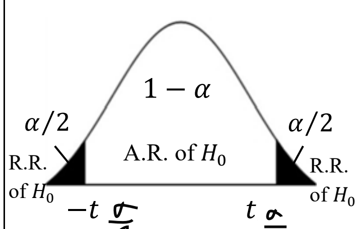
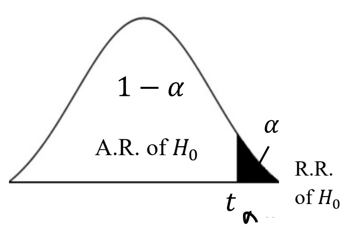
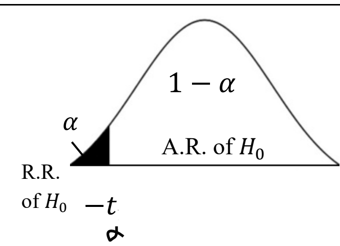
  P(Type II error) = P(Accepting H_0 | H_0 is false) = β
- The test statistic has the following form:

$$\text{Test Statistic} = \frac{\text{estimate} - \text{hypotheized parameter}}{\text{standard error of the estimate}}$$

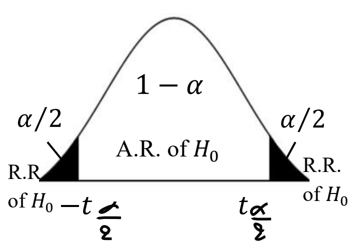
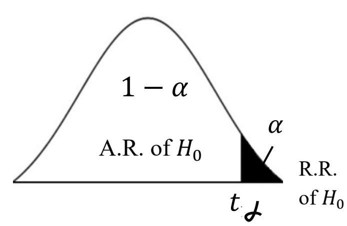
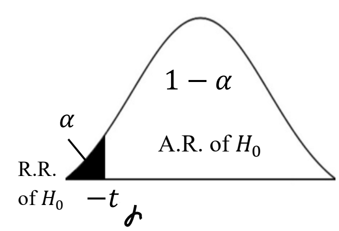
1. Hypotheses Testing for the population Mean (μ):

Hypotheses	$H_0: \mu = \mu_0$ vs $H_A: \mu \neq \mu_0$	$H_0: \mu = \mu_0$ vs $H_A: \mu > \mu_0$	$H_0: \mu = \mu_0$ vs $H_A: \mu < \mu_0$
Assumptions:	First Case: σ^2 is known; Normal or Non-Normal Distribution		
Test Statistic (T.S.)	$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$		
Rejection Region (R.R.) & Acceptance Region (A.R.) of H_0			
Critical Value	$-Z_{1-\frac{\alpha}{2}}$ and $Z_{1-\frac{\alpha}{2}}$	$Z_{1-\alpha}$	$-Z_{1-\alpha}$
Decision	We reject H_0 (and accept H_A) at the significance level α if:		
	$Z < -Z_{1-\frac{\alpha}{2}}$ or $Z > Z_{1-\frac{\alpha}{2}}$	$Z > Z_{1-\alpha}$	$Z < -Z_{1-\alpha}$
Assumptions:	Second Case: σ^2 is unknown; Normal Distribution		
Test Statistic (T.S.)	$T = \frac{\bar{X}-\mu_0}{S/\sqrt{n}} \sim t_{(n-1)}$; $\mathrm{d.f} = v = n-1$		
Rejection Region (R.R.) & Acceptance Region (A.R.) of H_0			
Critical Value	$-t_{\frac{\alpha}{2}}$ and $t_{\frac{\alpha}{2}}$	t_{α}	$-t_{\alpha}$
Decision	We reject H_0 (and accept H_A) at the significance level α if:		
	$T < -t_{\frac{\alpha}{2}}$ or $T > t_{\frac{\alpha}{2}}$	$T > t_{\alpha}$	$T < -t_{\alpha}$

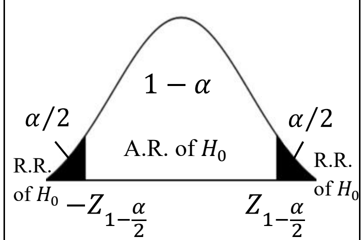
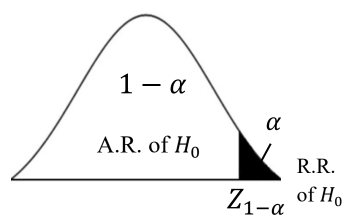
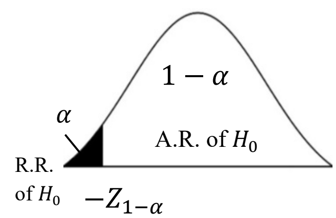
2. Hypotheses Testing for the Difference Between Two Population Means $(\mu_1 - \mu_2)$(Independent Populations):

Hypotheses	$H_0: \mu_1 = \mu_2$ vs $H_A: \mu_1 \neq \mu_2$	$H_0: \mu_1 = \mu_2$ vs $H_A: \mu_1 > \mu_2$	$H_0: \mu_1 = \mu_2$ vs $H_A: \mu_1 < \mu_2$
Assumptions:	First Case: σ_1^2 and σ_2^2 are known		
Test Statistic (T.S.)	$Z = \dfrac{\bar{X}_1 - \bar{X}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \sim N(0,1)$		
Rejection Region (R.R.) & Acceptance Region (A.R.) of H_0			
Critical Value	$-Z_{1-\frac{\alpha}{2}}$ and $Z_{1-\frac{\alpha}{2}}$	$Z_{1-\alpha}$	$-Z_{1-\alpha}$
Decision	We reject H_0 (and accept H_A) at the significance level α if:		
	$Z < -Z_{1-\frac{\alpha}{2}}$ or $Z > Z_{1-\frac{\alpha}{2}}$	$Z > Z_{1-\alpha}$	$Z < -Z_{1-\alpha}$
Two Sample Pooled t-Test Assumptions:	Second Case: σ_1^2 and σ_2^2 are unknown but equal $(\sigma_1^2 = \sigma_2^2 = \sigma^2)$		
Test Statistic (T.S.)	$T = \dfrac{\bar{X}_1 - \bar{X}_2}{\sqrt{\dfrac{S_p^2}{n_1} + \dfrac{S_p^2}{n_2}}} \sim t(n_1 + n_2 - 2), \quad df = v = n_1 + n_2 - 2$ $S_p^2 = \dfrac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$		
Rejection Region (R.R.) & Acceptance Region (A.R.) of H_0			
Critical Value	$-t_{\alpha/2}$ and $t_{\frac{\alpha}{2}}$	t_α	$-t_\alpha$
Decision	We reject H_0 (and accept H_A) at the significance level α if:		
	$T < -t_{\alpha/2}$ or $T > t_{\alpha/2}$	$T > t_\alpha$	$T < -t_\alpha$

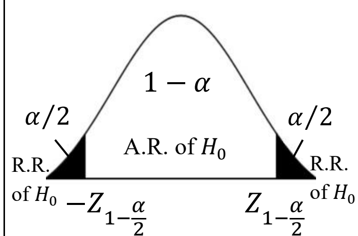
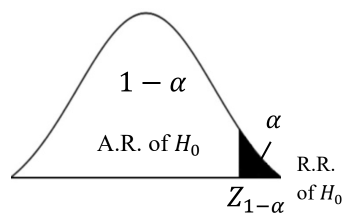
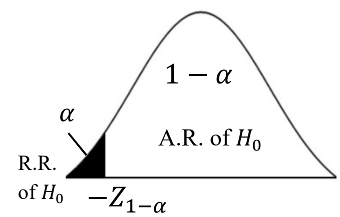
3. Confidence Interval and Hypotheses Testing for the Difference Between Two Population Means ($\boldsymbol{\mu_1 - \mu_2 = \mu_D}$) for Dependent (Related) Populations: Paired t-Test:

Calculate the Quantities	• The differences (D-observations): $D_i = X_i - Y_i\ ,\ \ i = 1,2,\dots,n$ • Sample Mean of the D-observations: $\overline{D} = \frac{\sum_{i=1}^{n} D_i}{n}$ • Sample Variance of the D-observations: $S_D^2 = \frac{\sum_{i=1}^{n}(D_i - \overline{D})^2}{n-1}$ • Sample Standard Deviation of the D-observations: $S_D = \sqrt{S_D^2}$		
Hypotheses Testing for $\mu_D = \mu_1 - \mu_2$			
Hypotheses	$H_0: \mu_1 = \mu_2\ \ vs$ $H_A: \mu_1 \neq \mu_2$ or $H_0: \mu_D = 0\ \ vs$ $H_A: \mu_D \neq 0$	$H_0: \mu_1 \leq \mu_2\ \ vs$ $H_A: \mu_1 > \mu_2$ or $H_0: \mu_D \leq 0\ \ vs$ $H_A: \mu_D > 0$	$H_0: \mu_1 \geq \mu_2\ \ vs$ $H_A: \mu_1 < \mu_2$ or $H_0: \mu_D \geq 0\ \ vs$ $H_A: \mu_D < 0$
Test Statistic (T.S.)	$T = \frac{\overline{D}}{S_D/\sqrt{n}} \sim t(n-1)\ ,\quad df = v = n-1$		
Rejection Region (R.R.) & Acceptance Region (A.R.) of H_0			
Critical Value	$-t_{\frac{\alpha}{2}}$ and $t_{\frac{\alpha}{2}}$	t_α	$-t_\alpha$
Decision	We reject H_0 (and accept H_A) at the significance level α if:		
	$T < -t_{\frac{\alpha}{2}}$ or $T > t_{\frac{\alpha}{2}}$	$T > t_\alpha$	$T < -t_\alpha$

4. Hypotheses Testing for the Population Proportion (p):

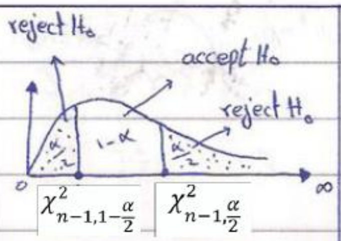
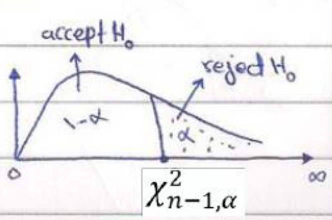
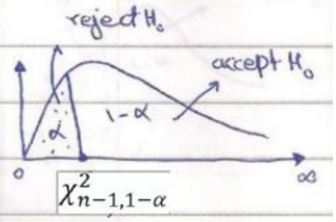
Hypotheses	$H_0: p = p_0 \ \ vs$ $H_A: p \neq p_0$	$H_0: p = p_0 \ \ vs$ $H_A: p > p$	$H_0: p = p_0 \ \ vs$ $H_A: p < p_0$
Test Statistic (T.S.)	$Z = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}} \sim N(0,1)\ , \qquad \hat{p} = \dfrac{X}{n}$		
Rejection Region (R.R.) & Acceptance Region (A.R.) of H_0			
Critical Value	$-Z_{1-\frac{\alpha}{2}}$ and $Z_{1-\frac{\alpha}{2}}$	$Z_{1-\alpha}$	$-Z_{1-\alpha}$
Decision	We reject H_0 (and accept H_A) at the significance level α if:		
	$Z < -Z_{1-\frac{\alpha}{2}}$ or $Z > Z_{1-\frac{\alpha}{2}}$	$Z > Z_{1-\alpha}$	$Z < -Z_{1-\alpha}$

5. Hypotheses Testing for the Difference Between Two Population Proportions $(p_1 - p_2)$:

Hypotheses	$H_0: p_1 = p_2 \ \ vs$ $H_A: p_1 \neq p_2$	$H_0: p_1 = p_2 \ \ vs$ $H_A: p_1 > p_2$	$H_0: p_1 = p_2 \ \ vs$ $H_A: p_1 < p_2$
Test Statistic (T.S.)	$Z = \dfrac{\hat{p}_1 - \hat{p}_2}{\sqrt{\dfrac{\bar{p}(1-\bar{p})}{n_1} + \dfrac{\bar{p}(1-\bar{p})}{n_2}}} \sim N(0,1)$ $\hat{p}_1 = \dfrac{X_1}{n_1}, \quad \hat{p}_2 = \dfrac{X_2}{n_2}, \quad \bar{p} = \dfrac{X_1 + X_2}{n_1 + n_2}$ $= \hat{p} =$ pooled estimate of common proportion		
Rejection Region (R.R.) & Acceptance Region (A.R.) of H_0			
Critical Value	$-Z_{1-\frac{\alpha}{2}}$ and $Z_{1-\frac{\alpha}{2}}$	$Z_{1-\alpha}$	$-Z_{1-\alpha}$
Decision	We reject H_0 (and accept H_A) at the significance level α if:		
	$Z < -Z_{1-\frac{\alpha}{2}}$ or $Z > Z_{1-\frac{\alpha}{2}}$	$Z > Z_{1-\alpha}$	$Z < -Z_{1-\alpha}$

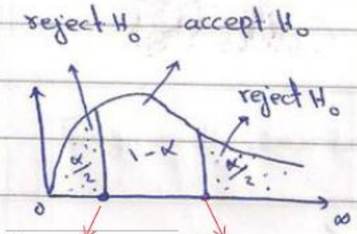
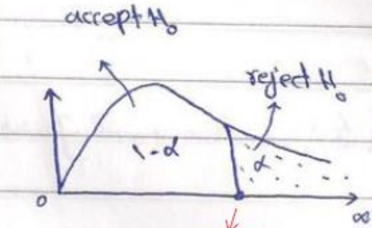
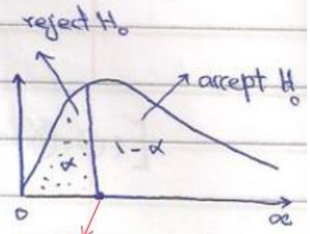
6. Hypotheses Testing for the population variance σ^2:

<u>Assumptions</u>: the distribution of the population being sampled is **normal**.

① H_0 and H_1	$H_0: \sigma^2 = \sigma_0^2$ Vs $H_1: \sigma^2 \neq \sigma_0^2$	$H_0: \sigma^2 = \sigma_0^2$ Vs $H_1: \sigma^2 > \sigma_0^2$	$H_0: \sigma^2 = \sigma_0^2$ Vs $H_1: \sigma^2 < \sigma_0^2$
② test statistic	$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2} \sim \chi^2_{n-1}$		
③ R.R and A.R of H_0, and table value			
④ decision	accept H_0 if $\chi^2_{n-1,1-\frac{\alpha}{2}} < \chi^2 < \chi^2_{n-1,\frac{\alpha}{2}}$ otherwise reject H_0	accept H_0 if $\chi^2 < \chi^2_{n-1,\alpha}$ otherwise reject H_0	accept H_0 if $\chi^2_{n-1,1-\alpha} < \chi^2$ otherwise reject H_0

7. Hypotheses Testing for the population variance σ_1^2 / σ_2^2 :

<u>Assumptions</u>: If two independent samples of size n_1 and n_2 are drawn at random from two **normal populations**.

① H_0 and H_1	$H_0: \sigma_1^2=\sigma_2^2 \longrightarrow \frac{\sigma_1^2}{\sigma_2^2}=1$ vs $H_1: \sigma_1^2\neq\sigma_2^2 \longrightarrow \frac{\sigma_1^2}{\sigma_2^2}\neq 1$	$H_0: \sigma_1^2=\sigma_2^2 \longrightarrow \frac{\sigma_1^2}{\sigma_2^2}=1$ vs $H_1: \sigma_1^2>\sigma_2^2 \longrightarrow \frac{\sigma_1^2}{\sigma_2^2}>1$	$H_0: \sigma_1^2=\sigma_2^2 \longrightarrow \frac{\sigma_1^2}{\sigma_2^2}=1$ vs $H_1: \sigma_1^2<\sigma_2^2 \longrightarrow \frac{\sigma_1^2}{\sigma_2^2}<1$
② test statistic	$F=\frac{S_1^2}{S_2^2} \sim F_{n_1-1,n_2-1}$		
③ R.R and A.R of H_0, and table value	 $F_{n_1-1,n_2-1,1-\frac{\alpha}{2}}$ $F_{n_1-1,n_2-1,\frac{\alpha}{2}}$	 $F_{n_1-1,n_2-1,\alpha}$	 $F_{n_1-1,n_2-1,1-\alpha}$
④ decision	accept H_0 if $F_{n_1-1,n_2-1,1-\frac{\alpha}{2}} < F < F_{n_1-1,n_2-1,\frac{\alpha}{2}}$ otherwise reject H_0	accept H_0 if $F < F_{n_1-1,n_2-1,\alpha}$ otherwise reject H_0	accept H_0 if $F_{n_1-1,n_2-1,1-\alpha} < F$ otherwise reject H_0

Note that:

1) Critical Region is Rejection Region (R.R.)

2) If we accept H0 $\cong$ fail to reject H0 $\cong$ reject H1 $\cong$ fail to accept H1

$\cong$ The (Z or T or χ^2 or F)-statistic is <u>not</u> significant at the α level.

If we reject H0 $\cong$ fail to accept H0 $\cong$ accept H1 $\cong$ fail to reject H1

$\cong$ The (Z or T or χ^2 or F)-statistic is significant at the α level.