Binomial	$P(X=x) = \binom{n}{x}p^x(1-p)^{n-x}$	E(X) = np, Var(X) = np(1-p)
Poisson	$P(X=x) = \frac{\lambda^x e^{-\lambda}}{x!}$	E(X) = λ, Var(X) = λ
Normal	$f(x) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)e^{-\left(\frac{1}{2}\right)\left(\frac{x-\mu}{\sigma}\right)^2}$	E(X) = μ, Var(X) = σ^2
Exponential	$f(x) = \lambda e^{-\lambda x}$ for $x \geq 0$	E(X) = $\frac{1}{\lambda}$, Var(X) = $\frac{1}{\lambda^2}$
Uniform	$f(x) = \frac{1}{b-a}$ for $a \leq x \leq b$	E(X) = $\frac{a+b}{2}$, Var(X) = $\frac{(b-a)^2}{12}$
Bernoulli	$P(X=1) = p$ $P(X=0) = 1-p$	E(X) = p, Var(X) = $p(1-p)$
Geometric	$P(X=x) = (1-p)^{x-1}p$	E(X) = $\frac{1}{p}$, Var(X) = $\frac{1-p}{p^2}$
Negative Binomial	$P(X=x) = \binom{x-1}{r-1}p^r(1-p)^{x-r}$	E(X) = $\frac{r}{p}$, Var(X) = $\frac{r(1-p)}{p^2}$
Hypergeometric	$P(X=x) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}$	E(X) = $n\left(\frac{K}{N}\right)$, Var(X) = $n\left(\frac{K}{N}\right)\left(1-\frac{K}{N}\right)\left(\frac{N-n}{N-1}\right)$

- MME: equate sample mean to population mean
- 1st moment: $\overline{X} = E(X)$
- 2nd moment: $\left(\frac{1}{n}\right)\sum_{i=1}^{n}(X_i^2) = E(X^2)$
- MLE: maximise the likelihood function(for n independent observations: $L(\theta) = \prod_{i=1}^{n} f(X_i;\theta)$). Also show the maxima using the 2nd derivative test.

	H_0 true	H_0 false
Fail to reject H_0 (i.e. -ve test)	Correct $1-\alpha$	Type II error
Reject H_0 (i.e. +ve test)	Type I error α	Correct Power of Test

1. HT for population mean

Use when: Testing if a population mean differs from a hypothesized value (use z if σ known, t if unknown); e.g. has the avg rate gone up (given μ) - this will b a one tailed test. e.g. one-taild test

- $H_0 : \mu = \mu_0 \mid H_1 : \mu > \mu_0$
- calculate sample mean $\overline{x}$ and sample standard deviation s_x
- calculate test statistic: $t_{\text{calc}} = \frac{\overline{x}-\mu_0}{\frac{s}{\sqrt{n}}}$ with df = n-1 (or use z if σ known)
- determine critical value from t-table (or z-table) at significance level α
- compare t_{calc} with t_{crit}
- If $t_{\text{calc}} > t_{\text{crit}}$ (for right-tailed test), reject H_0; If $t_{\text{calc}} < -t_{\text{crit}}$ (for left-tailed test), reject H_0

Comparison of one-tailed tests:

- **Right-tailed** ($H_1 : \mu > \mu_0$): Used when testing for an **increase** (e.g., "has performance improved?"). Reject if $t_{\text{calc}} > t_{\text{crit}}$ (positive critical value)
- **Left-tailed** ($H_1 : \mu < \mu_0$): Used when testing for a **decrease** (e.g., "has cost reduced?"). Reject if $t_{\text{calc}} < -t_{\text{crit}}$ (negative critical value)

2. HT for population proportion

Use when: Testing if a population proportion differs from a hypothesized value; e.g., testing if the proportion of car owners in a region differs from a claimed percentage.

- State hypotheses: $H_0 : P = p_0$ (population proportion equals hypothesized value); H_1: (choose: $P \neq p_0$ or $P > p_0$ or $P < p_0$)
- Calculate sample proportion: $\hat{p} = \frac{x}{n}$ where x is number of successes and n is sample size
- Check conditions: $np_0 \geq 10$ and $n(1-p_0) \geq 10$ for normal approximation validity
- Specify significance level α (typically 0.05) and calculate test statistic: $z_{\text{calc}} = \frac{\hat{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$
- Determine critical value(s) from z-table: Two-tailed: $\pm z_{\text{crit}}$ (e.g., ± 1.96 for $\alpha = 0.05$); Right-tailed: z_{crit} (e.g., 1.645 for $\alpha = 0.05$); Left-tailed: $-z_{\text{crit}}$ (e.g., -1.645 for $\alpha = 0.05$)
- Make decision: Two-tailed: If $|z_{\text{calc}}| > z_{\text{crit}}$, reject H_0; Right-tailed: If $z_{\text{calc}} > z_{\text{crit}}$, reject H_0; Left-tailed: If $z_{\text{calc}} < -z_{\text{crit}}$, reject H_0

3. HT for population variance

Use when: Testing if a population variance differs from a hypothesized value, or when you need to verify assumptions about data variability (e.g., quality control - has the variance in product weights changed?).

- $H_0 : \sigma^2 = \sigma_0^2$ (population variance equals hypothesized value); H_1: (choose: $\sigma^2 \neq \sigma_0^2$ or $\sigma^2 > \sigma_0^2$ or $\sigma^2 < \sigma_0^2$)
- Calculate sample variance: $s^2 = \left(\frac{1}{n-1}\right)\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$
- Specify significance level α (typically 0.05) and calculate test statistic: $\chi^2_{\text{calc}} = \frac{(n-1)s^2}{\sigma_0^2}$ with df = n-1
- Determine critical value(s) from χ^2 table: Two-tailed: χ^2_{lower} and χ^2_{upper} (e.g., at $\frac{\alpha}{2}$ and $1-\frac{\alpha}{2}$); Right-tailed: χ^2_{crit} at $1-\alpha$; Left-tailed: χ^2_{crit} at α
- Make decision: Two-tailed: If $\chi^2_{\text{calc}} < \chi^2_{\text{lower}}$ or $\chi^2_{\text{calc}} > \chi^2_{\text{upper}}$, reject H_0; Right-tailed: If $\chi^2_{\text{calc}} > \chi^2_{\text{crit}}$, reject H_0; Left-tailed: If $\chi^2_{\text{calc}} < \chi^2_{\text{crit}}$, reject H_0

4. 2 sample test + paired t test

Use when: Testing if the mean difference between paired observations is zero; e.g., before-after measurements on the same subjects (drug effectiveness, training impact), matched pairs (father-son heights, twin studies).

- Calculate differences: $d_i = x_i - y_i$ for each paired observation
- Calculate sample mean of differences: $\overline{d} = \frac{\sum_{i=1}^{n} d_i}{n}$
- Calculate sample standard deviation of differences: $s_d = \sqrt{\frac{\sum_{i=1}^{n}\left(d_i-\overline{d}\right)^2}{n-1}}$
- State hypotheses: $H_0 : \mu_d = 0$ (no difference/drug has no effect); H_1: (choose: $\mu_d \neq 0$ or $\mu_d > 0$ (i.e. drug improves score) or $\mu_d < 0$)
- Specify significance level α (typically 0.05) and calculate test statistic: $t_{\text{calc}} = \frac{\overline{d}-\mu_{d_0}}{\frac{s_d}{\sqrt{n}}}$ with df = n-1
- Determine critical value from t-table at chosen α and df
- Make decision: Two-tailed: If $|t_{\text{calc}}| > t_{\text{crit}}$, reject H_0; Right-tailed: If $t_{\text{calc}} > t_{\text{crit}}$, reject H_0; Left-tailed: If $t_{\text{calc}} < -t_{\text{crit}}$, reject H_0

5. 2 sample test + comparing mean

Use when: Testing if the means of two independent populations differ; e.g., comparing average scores between two groups, testing if mean salaries differ between two departments, comparing average heights between two different populations; e.g. TGA and TGB

(pooled variance - when σ^2 known to be equal)

- State hypotheses: $H_0 : \mu_1 = \mu_2$ or $H_0 : \mu_1 - \mu_2 = 0$; H_1 : (choose: $\mu_1 \neq \mu_2$ or $\mu_1 > \mu_2$ or $\mu_1 < \mu_2$)
- Calculate pooled variance: $s_p^2 = \frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}$
- Calculate test statistic: $t_{\text{calc}} = \frac{\overline{x_1}-\overline{x_2}}{\sqrt{s_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$ with df = $n_1 + n_2 - 2$
- Determine critical value from t-table at significance level α and df
- Make decision: Two-tailed: If $|t_{\text{calc}}| > t_{\text{crit}}$, reject H_0; Right-tailed: If $t_{\text{calc}} > t_{\text{crit}}$, reject H_0; Left-tailed: If $t_{\text{calc}} < -t_{\text{crit}}$, reject H_0

(pooled variance) CI = $\left((\overline{x_1}-\overline{x_2}) \pm t_{\text{crit}} * \sqrt{s_p^2 * \left(\frac{1}{n_1}+\frac{1}{n_2}\right)}\right)$

(unpooled variance - Welch's t-test - when σ^2 not known to be equal)

- State hypotheses: $H_0 : \mu_1 = \mu_2$; H_1 : (choose: $\mu_1 \neq \mu_2$ or $\mu_1 > \mu_2$ or $\mu_1 < \mu_2$)
- Calculate test statistic: $t_{\text{calc}} = \frac{\overline{x_1}-\overline{x_2}}{\sqrt{\left(\frac{s_1^2}{n_1}\right)+\left(\frac{s_2^2}{n_2}\right)}}$
- Calculate degrees of freedom: df = $\frac{\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1-1}+\frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2-1}}$
- Determine critical value from t-table at significance level α and calculated df
- Make decision: Two-tailed: If $|t_{\text{calc}}| > t_{\text{crit}}$, reject H_0; Right-tailed: If $t_{\text{calc}} > t_{\text{crit}}$, reject H_0; Left-tailed: If $t_{\text{calc}} < -t_{\text{crit}}$, reject H_0

(unpooled variance) CI = $\left((\overline{x_1}-\overline{x_2}) \pm t_{\text{crit}} * \sqrt{\left(\frac{s_1^2}{n_1}\right)+\left(\frac{s_2^2}{n_2}\right)}\right)$ and df calculation for unpooled case: df = $\frac{\left(\left(\frac{s_1^2}{n_1}\right)+\left(\frac{s_2^2}{n_2}\right)\right)^2}{\left(\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1-1}\right)+\left(\frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2-1}\right)}$

6. 2 sample test + comparing proportion

Use when: Testing if proportions differ between two independent groups; e.g., comparing preference rates, success rates, or occurrence rates between two populations; e.g. PP non-smoking rooms preference | e.g. the cardiac arrests during day and night number of deaths question | e.g. elem and high school teacher preference.

Decision Rule:

- Use **pooled proportion** for hypothesis testing (assumes $p_1 = p_2$ under H_0)
- Use **unpooled proportion** for confidence intervals (no assumption about equality)

(unpooled proportion - for confidence intervals) CI = $(\hat{p_1} - \hat{p_2}) \pm z_{\text{crit}} * \sqrt{\frac{\hat{p_1}(1-\hat{p_1})}{n_1} + \frac{\hat{p_2}(1-\hat{p_2})}{n_2}}$ (pooled proportion - for hypothesis testing)

- State hypotheses: $H_0 : p_1 = p_2$ or $H_0 : p_1 - p_2 = 0$; H_1 : (choose: $p_1 \neq p_2$ or $p_1 > p_2$ or $p_1 < p_2$)
- Calculate sample proportions: $\hat{p_1} = \frac{x_1}{n_1}$ and $\hat{p_2} = \frac{x_2}{n_2}$ and pooled proportion: $\hat{p} = \frac{x_1+x_2}{n_1+n_2}$

- Calculate test statistic: $z_{\text{calc}} = \frac{\widehat{p}_1 - \widehat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$
- Determine critical value from z-table based on significance level α and test type (one-tailed or two-tailed)
- Make decision: If $|z_{\text{calc}}| > z_{\text{crit}}$ (two-tailed) or $z_{\text{calc}} > z_{\text{crit}}$ (right-tailed) or $z_{\text{calc}} < -z_{\text{crit}}$ (left-tailed), reject H_0.

7. 2 sample test + comparing variance

Use when: Testing if the variances of two independent populations differ; e.g., comparing variability in quality between two manufacturing processes, testing if variance in test scores differs between two teaching methods.

- State hypotheses: $H_0 : \sigma_1^2 = \sigma_2^2$ or $H_0 : \frac{\sigma_1^2}{\sigma_2^2} = 1$; H_1 : (choose: $\sigma_1^2 \neq \sigma_2^2$ or $\sigma_1^2 > \sigma_2^2$ or $\sigma_1^2 < \sigma_2^2$)
- Calculate sample variances: s_1^2 and s_2^2 from each sample
- Specify significance level α (typically 0.05)
- Calculate test statistic: $F_{\text{calc}} = \frac{s_1^2}{s_2^2}$ (always put larger variance in numerator for one-tailed test, or follow hypothesis direction for two-tailed)
- Determine degrees of freedom: $\text{df}_1 = n_1 - 1$ (numerator) and $\text{df}_2 = n_2 - 1$ (denominator)
- Determine critical value(s) from F-distribution table: Two-tailed: F_{lower} at $\frac{\alpha}{2}$ and F_{upper} at $1 - \frac{\alpha}{2}$; Right-tailed: F_{crit} at $1 - \alpha$; Left-tailed: F_{crit} at α
- Make decision: Two-tailed: If $F_{\text{calc}} < F_{\text{lower}}$ or $F_{\text{calc}} > F_{\text{upper}}$, reject H_0; Right-tailed: If $F_{\text{calc}} > F_{\text{crit}}$, reject H_0; Left-tailed: If $F_{\text{calc}} < \frac{1}{F_{\text{crit}}}$, reject H_0

Note: F-test is sensitive to non-normality. Consider Levene's test or Bartlett's test as alternatives when normality is questionable.

8. ANOVA

Use when: Testing if means of three or more independent groups differ; e.g., comparing average test scores across multiple teaching methods, testing if mean sales differ across different regions. **Note:** ANOVA assumes normality, independence, and equal variances across groups. If assumptions are violated, consider Kruskal-Wallis test.

- **State hypotheses:** $H_0 : \mu_1 = \mu_2 = \mu_3 = \cdots = \mu_k$ (all group means are equal); H_1 : at least one group mean differs from the others
- Calculate group means $\overline{X}_i$ for each of the k groups and overall mean $\overline{X}$
- Calculate Sum of Squares Between groups: $\text{SSB} = \sum_{i=1}^{k} n_i\left(\overline{X}_i - \overline{X}\right)^2$
- Calculate Sum of Squares Within groups: $\text{SSW} = \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(X_{ij} - \overline{X}_i\right)^2$
- Calculate Mean Squares: $\text{MSB} = \frac{\text{SSB}}{k-1}$ and $\text{MSW} = \frac{\text{SSW}}{N-k}$ where $N = \sum n_i$
- Calculate test statistic: $F_{\text{calc}} = \frac{\text{MSB}}{\text{MSW}}$ and determine degrees of freedom: $\text{df}_1 = k - 1$ (between groups) and $\text{df}_2 = N - k$ (within groups)
- Specify significance level α (typically 0.05) and find critical value F_{crit} from F-table
- Make decision: If $F_{\text{calc}} > F_{\text{crit}}$, reject H_0 (at least one group mean differs)

ANOVA Table:

Source	df	Sum of Squares	Mean Square	F-statistic
Between Groups	$k-1$	SSB	$\text{MSB} = \frac{\text{SSB}}{k-1}$	$F = \frac{\text{MSB}}{\text{MSW}}$
Within Groups	$N-k$	SSW	$\text{MSW} = \frac{\text{SSW}}{N-k}$	
Total	$N-1$	SST		

where $\text{SST} = \text{SSB} + \text{SSW}$

9. Bartlett's Test for Homogeneity of Variances

Use when: Testing if variances are equal across multiple groups (k ≥ 2); e.g., comparing variability in yields across different fertilizer treatments (additives A, B, C). **Note:** Bartlett's test is sensitive to non-normality. Consider Levene's test as a more robust alternative.

- **State hypotheses:** $H_0 : \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_k^2$ (all population variances are equal); H_1 : at least one variance differs.
- **Calculate pooled variance:** $S_p^2 = \frac{\sum_{i=1}^{k}(n_i-1)s_i^2}{N-k}$ where $N = \sum_{i=1}^{k} n_i$ (total sample size)
- **Calculate test statistic:** $\chi^2_{\text{calc}} = \frac{(N-k)\ln(S_p^2) - \sum_{i=1}^{k}(n_i-1)\ln(s_i^2)}{C}$ where $C = 1 + \left(\frac{1}{3(k-1)}\right)\left(\sum_{i=1}^{k}\left(\frac{1}{n_i-1}\right) - \left(\frac{1}{N-k}\right)\right)$ and **Determine degrees of freedom:** $\text{df} = k - 1$
- **Specify significance level** α (typically 0.05) and find critical value χ^2_{crit} from χ^2 table
- **Make decision:** If $\chi^2_{\text{calc}} > \chi^2_{\text{crit}}$, reject H_0 (variances are not equal)

10. 2 way ANOVA

Use when: Testing the effects of two categorical independent variables (factors) on a continuous dependent variable, including their interaction; e.g., studying effect of fertilizer type AND irrigation method on crop yield, or teaching method AND class size on test scores. **Assumptions:** Independent observations + Normally distributed residuals + Homogeneity of variance across groups + Balanced design (equal sample sizes per cell) preferred

- **Hypotheses:** Factor A: $H_0 : \alpha_1 = \alpha_2 = \cdots = \alpha_r = 0$ (no effect of factor A); H_1 : at least one $\alpha_i \neq 0$ | Factor B: $H_0 : \beta_1 = \beta_2 = \cdots = \beta_c = 0$ (no effect of factor B); H_1 : at least one $\beta_j \neq 0$ | Interaction: H_0 : no interaction between A and B; H_1 : interaction exists
- **Calculate cell means** $\overline{X}_{ij}$, row means $\overline{X}_{i\cdot}$, column means $\overline{X}_{\cdot j}$, and grand mean $\overline{X}$
- **Calculate Sum of Squares:**
  - Total: $\text{SST} = \sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k=1}^{n}\left(X_{ijk} - \overline{X}\right)^2$
  - Factor A: $\text{SS}_A = cn\sum_{i=1}^{r}\left(\overline{X}_{i\cdot} - \overline{X}\right)^2$
  - Factor B: $\text{SS}_B = rn\sum_{j=1}^{c}\left(\overline{X}_{\cdot j} - \overline{X}\right)^2$
  - Interaction: $\text{SS}_{AB} = n\sum_{i=1}^{r}\sum_{j=1}^{c}\left(\overline{X}_{ij} - \overline{X}_{i\cdot} - \overline{X}_{\cdot j} + \overline{X}\right)^2$
  - Error: $\text{SSE} = \sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k=1}^{n}\left(X_{ijk} - \overline{X}_{ij}\right)^2$
- **Calculate Mean Squares:**
  - $\text{MS}_A = \frac{\text{SS}_A}{r-1}$
  - $\text{MS}_B = \frac{\text{SS}_B}{c-1}$
  - $\text{MS}_{AB} = \frac{\text{SS}_{AB}}{(r-1)(c-1)}$
  - $\text{MSE} = \frac{\text{SSE}}{rc(n-1)}$
- **Calculate F-statistics:**
  - Factor A: $F_A = \frac{\text{MS}_A}{\text{MSE}}$ with df = $(r-1, rc(n-1))$
  - Factor B: $F_B = \frac{\text{MS}_B}{\text{MSE}}$ with df = $(c-1, rc(n-1))$
  - Interaction: $F_{AB} = \frac{\text{MS}_{AB}}{\text{MSE}}$ with df = $((r-1)(c-1), rc(n-1))$
- **Specify significance level** α (typically 0.05) and **Find critical values** from F-table for each test
- **Make decisions:** For each test, if $F_{\text{calc}} > F_{\text{crit}}$, reject H_0

Two-Way ANOVA Table:

Source	df	Sum of Squares	Mean Square	F-statistic
Factor A	$r-1$	SS_A	MS_A	$F_A = \frac{\text{MS}_A}{\text{MSE}}$
Factor B	$c-1$	SS_B	MS_B	$F_B = \frac{\text{MS}_B}{\text{MSE}}$
Interaction A×B	$(r-1)(c-1)$	SS_{AB}	MS_{AB}	$F_{AB} = \frac{\text{MS}_{AB}}{\text{MSE}}$
Error	$rc(n-1)$	SSE	MSE	
Total	$rcn-1$	SST		

where r = number of levels of factor A, c = number of levels of factor B, n = number of observations per cell.

Note: If interaction is significant, interpret main effects with caution as the effect of one factor depends on the level of the other factor.

11. χ^2 test + goodness of fit

Use when: Testing if observed categorical data follows a specified theoretical distribution; e.g., testing if die rolls are fair, if color preferences match expected proportions, or if genotypes follow Mendelian ratios.

e.g. rock paper scissor

- State hypotheses: H_0 : the observed frequencies fit the expected frequencies (data follows the specified distribution); H_1 : they do not fit
- Calculate expected counts: $E_i = \text{Total} \times \text{Probability}_i$ for each category
- Check condition: Ensure all expected frequencies $E_i \geq 5$ for validity
- Specify significance level α (typically 0.05)
- Calculate test statistic: $\chi^2_{\text{calc}} = \sum_{i=1}^{c}\left(\frac{(O_i - E_i)^2}{E_i}\right)$ where O_i are observed frequencies
- Determine degrees of freedom: $\text{df} = c - 1$ where c is the number of categories
- Find critical value χ^2_{crit} from χ^2 table at chosen α and df
- Make decision: If $\chi^2_{\text{calc}} > \chi^2_{\text{crit}}$, reject H_0 (observed data does not fit expected distribution). If $\chi^2_{\text{calc}} \leq \chi^2_{\text{crit}}$, fail to reject H_0 (insufficient evidence against the expected distribution)

12. χ^2 test + test of independence

e.g. illegal piracy law and attitude | e.g. accident and seatbelt usage

- H_0 : the two variables are independent (no association); H_1 : the two variables are not independent (there is an association)
- **Create contingency table** and calculate totals for rows, columns, and grand total N
- **Calculate expected frequencies:** $E_{\text{ij}} = \frac{\text{row}_i \text{ total} \times \text{column}_j \text{ total}}{N}$
- **Calculate test statistic:** $\chi^2_{\text{calc}} = \sum_{i,j}\frac{(O_{\text{ij}} - E_{\text{ij}})^2}{E_{\text{ij}}}$ where O_{ij} are observed frequencies and E_{ij} are expected frequencies
- **Determine degrees of freedom:** $\text{df} = (r-1)(c-1)$ where r = number of rows and c = number of columns
- **Specify significance level** α (typically 0.05) and **Find critical value** χ^2_{crit} from χ^2 table at chosen α and df
- **Make decision:** If $\chi^2_{\text{calc}} > \chi^2_{\text{crit}}$, reject H_0 (variables are dependent). If $\chi^2_{\text{calc}} \leq \chi^2_{\text{crit}}$, fail to reject H_0 (insufficient evidence of dependence)

Note: Ensure all expected frequencies $E_{\text{ij}} \geq 5$ for validity of chi-square approximation.

13. Wilcoxon Signed-Rank Test

Use when: Testing if the median of paired differences equals zero (or a hypothesized value) when normality assumption is violated; e.g., non-parametric alternative to paired t-test for before-after measurements, matched pairs when data is skewed or ordinal.

- **State hypotheses:** $H_0 : \tilde{\mu}_d = 0$ (median difference is zero); H_1 : (choose: $\tilde{\mu}_d \neq 0$ or $\tilde{\mu}_d > 0$ or $\tilde{\mu}_d < 0$)
- **Calculate differences:** $d_i = x_i - y_i$ for each paired observation
- **Rank absolute differences:** Ignore zero differences, rank $|d_i|$ from smallest to largest (assign average ranks for ties)
- **Assign signs:** Attach the original sign of each difference to its rank
- **Calculate test statistic:** $W = \sum$ of positive ranks (or sum of negative ranks - use smaller for two-tailed test)
- **Determine critical value:** For small samples ($n \leq 25$), use Wilcoxon signed-rank table; for large samples ($n > 25$), use normal approximation: $Z = \frac{W - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}}$
- **Specify significance level** α (typically 0.05)
- **Make decision:**
  - Small sample: If $W \leq W_{\text{crit}}$ from table, reject H_0
  - Large sample: If $|Z| > z_{\text{crit}}$ (two-tailed) or $Z > z_{\text{crit}}$ (right-tailed) or $Z < -z_{\text{crit}}$ (left-tailed), reject H_0

Note: More robust than paired t-test for non-normal data or outliers. Assumes symmetric distribution of differences around the median.

14. Wilcoxon Mann Whitney U Test

Use when: Testing if two independent samples come from populations with the same distribution (non-parametric alternative to two-sample t-test); e.g., comparing treatment effects when normality assumption is violated, or with ordinal data.

- **State hypotheses:** H_0 : the two populations have identical distributions (or equal medians); H_1 : the distributions differ (or medians differ)
- **Rank all observations:** Combine both samples, rank from smallest to largest (assign average ranks for ties)
- **Calculate rank sums:** $R_1 =$ sum of ranks for sample 1; $R_2 =$ sum of ranks for sample 2
- **Calculate U statistics:** $U_1 = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1$ and $U_2 = n_1 n_2 + \frac{n_2(n_2+1)}{2} - R_2$
- **Test statistic:** Use $U = \min(U_1, U_2)$ (or equivalently, can use U_1 or U_2 with appropriate critical values)
- **Determine critical value:**
  - Small samples ($n_1, n_2 \leq 20$): Use Mann-Whitney U table
  - Large samples: Use normal approximation: $Z = \frac{U - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}}$
- **Specify significance level** α (typically 0.05)
- **Make decision:**
  - Small sample: If $U \leq U_{\text{crit}}$ from table, reject H_0
  - Large sample: If $|Z| > z_{\text{crit}}$ (two-tailed) or $Z > z_{\text{crit}}$ (right-tailed) or $Z < -z_{\text{crit}}$ (left-tailed), reject H_0

Note: More robust than two-sample t-test for non-normal data, outliers, or ordinal data. Does not assume equal variances.

15. Kruskal-Wallis Test

Use when: Comparing medians across 3+ independent groups with non-normal data or ordinal scales (non-parametric alternative to one-way ANOVA); e.g., comparing satisfaction ratings across multiple brands, test scores across teaching methods when normality is violated; e.g. MM kitkat calorie question

- H_0 : the medians are equal across all groups (or distributions are identical); H_1 : not all medians are equal (at least one group differs)
- Combine all observations from all groups and rank them from smallest to largest (assign average ranks for ties)
- Sum the ranks for each group: $R_1, R_2, ..., R_k$
- Calculate test statistic: $H = \left(\frac{12}{N(N+1)}\right) \sum_{j=1}^{k} \left(\frac{R_j^2}{n_j}\right) - 3(N+1)$ where N is total sample size and n_j is size of group j
- Determine degrees of freedom: df $= k - 1$ where k is number of groups
- Find critical value from χ^2 table at chosen significance level α
- Decision: If $H_{\text{calc}} > \chi^2_{\text{crit}}$, reject H_0

Note: Use when ANOVA assumptions are violated (non-normal distributions, unequal variances) or with ordinal data. For large samples, H follows χ^2 distribution approximately.

On Regression

Total Variance (SST) = Variance Explained (SSR) + Variance Unexplained (SSE)

$\text{SST} = \sum_{i=1}^{n} \left(Y_i - \overline{Y}\right)^2$ $\text{SSR} = \sum_{i=1}^{n} \left(\hat{Y}_i - \overline{Y}\right)^2$ $\text{SSE} = \sum_{i=1}^{n} \left(Y_i - \hat{Y}_i\right)^2$
$\text{SST} = \text{SSR} + \text{SSE}$

Coefficient of Determination $R^2 = \frac{\text{SSR}}{\text{SST}} = 1 - \left(\frac{\text{SSE}}{\text{SST}}\right)$

F-distributed test statistic

$H_0 : \beta_1 = \beta_2 = \beta_3 = c... = \beta_k = 0$
H_1 : at least one $\beta_j \neq 0$
k is the number of independent variables
$\text{MSR} = \frac{\text{SSR}}{k}$
$\text{MSE} = \frac{\text{SSE}}{n-k-1}$ where $n - k - 1$ is the degrees of freedom
n is the number of observations and k is the number of independent variables

$F = \frac{\frac{\text{SSR}}{k}}{\frac{\text{SSE}}{n-k-1}} = \frac{\text{MSR}}{\text{MSE}}$

- The F-statistic in regression analysis is one-sided with the rejection region in the right tail of the F-distribution.
- If $F_{\text{calc}} > F_{\text{crit}}$, reject H_0.

Hypothesis test for individual regression coefficients

Test of slope coeffecient

$$t = \frac{\hat{b_1} - B_1}{s_{\hat{b_1}}}$$

This is t distributed with $n - k - 1$ degrees of freedom.

where $s_{\hat{b_1}}$ is the standard error of the slope coefficient.

$$s_{\hat{b_1}} = \sqrt{\frac{\text{MSE}}{\sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2}}$$

e.g. of linear regression b/w X and Y - test whether the slope coeffecient is different from 0 to confirm if there is a signifiant relationship between X and Y.

1. state hypothesis

$H_0 : \beta_1 = 0$ $H_1 : \beta_1 \neq 0$

2. Identify appropriate test statistic with $t = \frac{\hat{b_1} - B_1}{s_{\hat{b_1}}}$ with n - k - 1 = n - 2 degrees of freedom (with k = 1).
3. specify significance level α (say 0.05)
4. state the decision rule by determining the critical value from t-distribution table. In this case, critical t-values = ± 2.776. Reject H_0 if $t_{\text{calc}} < -2.776$ or $t_{\text{calc}} > 2.776$.
5. Calculate test statistic. Suppose slope coeffecient $\hat{b_1} = 1.25$, $\text{MSE} = 11.9688$ and variation of X = 122.640. Then, $s_e = \sqrt{\text{MSE}} = \sqrt{11.9688} = 3.46$. $s_{\hat{b_1}} = \frac{3.46}{\sqrt{122.640}} = 0.31$. $t = \frac{1.25 - 0}{0.31} = 4$
6. Make a decision. Since $t_{\text{calc}} = 4$ is greater than 2.776, we reject H_0 of zero slope and conclude that there is a significant linear relationship between X and Y.

Test of correlation

1. State the hypothesis. $H_0 : \rho = 0$ $H_1 : \rho \neq 0$
2. Identify appropriate test statistic with

$$t = r\sqrt{\frac{n-2}{1-r^2}}$$

with n - k - 1 degrees of freedom = n - 2 (with k = 1).

3. specify significance level α (say 0.05)
4. state the decision rule by determining the critical value from t-distribution table. In this case, critical t-values = ± 2.776. Reject H_0 if $t_{\text{calc}} < -2.776$ or $t_{\text{calc}} > 2.776$.
5. calculate the test statistic. correlation (r) is 0.8945 and n = 6. $t = 0.8945\frac{\sqrt{4}}{\sqrt{1-0.8945^2}} = 4$.
6. Make a decision. Since $t_{\text{calc}} = 4$ is greater than 2.776, we reject H_0 of zero correlation and conclude that there is significant evidence of correlation between X and Y.

e.g. test whether there is a positive slope or positive correlation. same steps as above. Except this is a one tailed test. So only check if $t_{\text{calc}} > t_{\text{crit}}$. $t_{\text{crit}} = +2.312$ for 5% level of significance with 4 degrees of freedom.

Tests of the intercept

$$t_{\text{intercept}} = \frac{\hat{b_0} - B_0}{s_{\hat{b_0}}}$$

where $s_{\hat{b_0}} = S_e\sqrt{\frac{1}{n} + \frac{\overline{X}^2}{\sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2}}$

On R^2 and Adjusted R^2 scores

R^2 (Coefficient of Determination)

Formula:

$$R^2 = \frac{\text{SSR}}{\text{SST}} = 1 - \left(\frac{\text{SSE}}{\text{SST}}\right)$$

where:
- $\text{SST} = \sum_{i=1}^{n} \left(Y_i - \overline{Y}\right)^2$ (Total Sum of Squares)
- $\text{SSR} = \sum_{i=1}^{n} \left(\hat{Y}_i - \overline{Y}\right)^2$ (Regression Sum of Squares - Explained Variation)
- $\text{SSE} = \sum_{i=1}^{n} \left(Y_i - \hat{Y}_i\right)^2$ (Error Sum of Squares - Unexplained Variation)

Meaning:
- Measures the proportion of total variation in the dependent variable (Y) that is explained by the independent variable(s)
- Range: $0 \leq R^2 \leq 1$

- $R^2 = 0.89$ means 89% of variation in Y is explained by the model
- Higher R^2 indicates better model fit
- R^2 always increases (or stays the same) when adding more predictors, even if they are not meaningful

Adjusted R^2

Formula: $R^2_{\text{adj}} = 1 - \left(\frac{n-1}{n-k-1}\right)(1 - R^2)$

or equivalently: $R^2_{\text{adj}} = 1 - \left(\frac{\frac{\text{SSE}}{n-k-1}}{\frac{\text{SST}}{n-1}}\right)$

where:
- n = number of observations
- k = number of independent variables (predictors)
- $n - k - 1$ = degrees of freedom for residuals

Meaning:
- Adjusts R^2 for the number of predictors in the model
- Penalizes for adding unnecessary variables
- Can decrease when adding predictors that don't improve the model enough
- Better for comparing models with different numbers of predictors
- Always: $R^2_{\text{adj}} \leq R^2$
- Use Adjusted R^2 to avoid overfitting and to compare models fairly

NOTE: relation b/w correlation and R^2 score; $r = \pm\sqrt{R^2 \text{ score}}$. The sign is same as that of the slope coeffecient.

G-test (Likelihood Ratio Test) for Logistic Regression

The G-test evaluates whether a logistic regression model with predictors significantly improves fit compared to a null model (intercept only).

Hypothesis

H_0: The predictor variable(s) provide no predictive value (null model is adequate)
H_1: At least one predictor significantly improves the model

Test Statistic

$$G = 2(\ln L_{\text{full}} - \ln L_{\text{null}}) = \text{Null Deviance} - \text{Residual Deviance}$$

where:
- $\ln L_{\text{full}}$ = log-likelihood of model with predictors
- $\ln L_{\text{null}}$ = log-likelihood of null model (intercept only)
- Deviance $= -2 \times \ln L$

Distribution: $G \sim \chi^2_k$ where k = number of predictors added

Decision Rule

- Calculate G_{calc} from model output
- Compare with χ^2_{crit} at significance level α with k degrees of freedom
- Reject H_0 if $G_{\text{calc}} > \chi^2_{\text{crit}}$ or p-value $< \alpha$

Example: Predicting Transmission Type from Weight (mtcars)

Model: Predict transmission type (`am`: 0=automatic, 1=manual) from car weight (`wt`)

Output Interpretation:

```
Null deviance:     43.230  on 31 degrees of freedom
Residual deviance: 19.176  on 30 degrees of freedom
```

Calculation:

1. Calculate G-statistic:

$$G = 43.230 - 19.176 = 24.054$$

2. Degrees of freedom: $k = 1$ (one predictor: weight)
3. Critical value: $\chi^2_{0.05^{(1)}} = 3.841$
4. Decision: $G_{\text{calc}} = 24.054 > 3.841 = \chi^2_{\text{crit}}$

   Therefore, **reject H_0**. Weight is a significant predictor of transmission type.
5. P-value: $p < 0.001$ (highly significant)

Interpretation:
- The model with weight fits significantly better than just predicting the overall proportion of manual vs automatic cars
- Adding weight reduces unexplained deviance by 24.054 units
- This improvement is statistically significant at any reasonable α level

When to Use G-test:
- Overall model significance testing (is the model better than guessing?)
- Comparing nested models (e.g., model with 3 predictors vs model with 1 predictor)
- More reliable than Wald test for small samples or extreme coefficients
- Particularly important in logistic regression where residuals are not normally distributed

Model Selection Criteria: AIC and BIC

AIC (Akaike Information Criterion)

Formula: $\text{AIC} = 2k - 2\ln(L)$

or for linear regression: $\text{AIC} = n\ln\left(\frac{\text{SSE}}{n}\right) + 2k$

where:
- k = number of parameters (including intercept): $k = p + 1$ where p is number of predictors
- L = maximum likelihood of the model
- n = number of observations
- SSE = sum of squared errors

Meaning & Interpretation:
- Measures model quality by balancing goodness of fit and model complexity
- Lower AIC indicates better model
- Penalizes for adding too many parameters (prevents overfitting)

Decision Rule: Compare AIC values across multiple candidate models and Choose the model with the **minimum AIC**

BIC (Bayesian Information Criterion)

Formula: $\text{BIC} = k\ln(n) - 2\ln(L)$

or for linear regression: $\text{BIC} = n\ln\left(\frac{\text{SSE}}{n}\right) + k\ln(n)$

Meaning:
- Similar to AIC but with stronger penalty for additional parameters
- Lower BIC indicates better model
- BIC penalizes complexity more heavily than AIC (especially for large n)
- Tends to select simpler models than AIC
- Choose model with **minimum BIC**

Comparison: R^2 vs Adjusted R^2 vs AIC vs BIC

Criterion	Goal	Penalty for Complexity	Best Value	Use Case
R^2	Maximize	None	Higher = Better	Overall fit assessment
Adjusted R^2	Maximize	Moderate	Higher = Better	Compare models with different # of predictors
AIC	Minimize	Moderate ($2k$)	Lower = Better	Model selection, n balance fit & complexity
BIC	Minimize	Strong ($k\ln(n)$)	Lower = Better	Model selection, n prefer simpler models

---

There is a difference between regression model and the fitted regression equation.

e.g. model

$$\text{Price} = \beta_0 + \beta_1\ \text{Area} + \beta_2\ \text{Bedrooms} + \varepsilon$$

fitted regression equation

$$\widehat{\text{Price}} = \widehat{\beta_0} + \widehat{\beta_1}\ \text{Area} + \widehat{\beta_2}\ \text{Bedrooms}$$

On deriving standard regression anova table from the sequential (type I) anova table

input

NOTE: this table is the output given by `anova(model)` in R

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
SqFt	1	82.778	82.778	27.8040	1.842e-05
Bedrooms	1	13.833	13.833	4.6464	0.04094
Bathrooms	1	14.021	14.021	4.7095	0.03970
factor(Mall)	1	37.137	37.137	12.4740	0.00163
Residuals	25	74.430	2.977		

output

Source	Df	Sum Sq	Mean Sq	F value
Regression	4	SSR = 82.778 + 13.833 + 14.021 + 37.137 = 147.769	MSR = SSR / DF = 147.769 / 4 = 36.942	$F_{\text{obs}} = \frac{\text{MSR}}{\text{MSE}} = \frac{36.942}{2.977} =$ 12.41
Residuals	25	SSE = 74.430	2.977	
Total	$4 + 25 = 29$	SST = SSR + SSE = 147.769 +		

		74.430 = 222.199		

some more notes on this regd hypothesis testing

5% significance level with 4 and 25 degrees of freedom
From F-table, critical value = 2.76
Since $F_{\text{obs}} = 12.41 > 2.76 = F_{\text{crit}}$, we reject H_0 and conclude that at least one of the regression coefficients is significantly different from zero.

on hypothesis testing for individual regression coefficients (say for SqFt)

$H_0 : \beta_1 = 0$ $H_1 : \beta_1 \neq 0$

n = 30, k = 4 (no. of independent variables), df = n - k - 1 = 30 - 4 - 1 = 25 this is a 2 tailed test. so critical t values = ± 2.060 (from t table for 25 df at 5% significance level). The given t-value in the table is 1.230. $1.230 < 2.060$, so we fail to reject H_0 and conclude that the regression coefficient for SqFt is not significantly different from zero. Alternatively, we can check the p-value = 0.0.23028 > 0.05, so we fail to reject H_0.

Another interesting question on ANOVA and regression

Logistic Regression

Generic Fitted Model for Logistic Regression

The logistic regression model predicts the probability that a binary outcome Y equals 1 given predictor variables $X_1, X_2, ..., X_p$.

Fitted Model:

$$\hat{P}(Y = 1 \mid X) = \frac{1}{1 + e^{-\left(\hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \cdots + \hat{\beta}_p X_p\right)}}$$

Or equivalently:

$$\hat{P}(Y = 1 \mid X) = \frac{e^{\hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \cdots + \hat{\beta}_p X_p}}{1 + e^{\hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \cdots + \hat{\beta}_p X_p}}$$

Fitted Logit (Log-Odds) Equation:

$$\text{logit}\left(\hat{P}\right) = \ln\left(\frac{\hat{P}}{1 - \hat{P}}\right) = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \cdots + \hat{\beta}_p X_p$$

Where:
- $\hat{P}(Y = 1 \mid X)$ is the estimated probability that $Y = 1$
- $\hat{\beta}_0$ is the estimated intercept
- $\hat{\beta}_1, \hat{\beta}_2, ..., \hat{\beta}_p$ are the estimated coefficients
- $X_1, X_2, ..., X_p$ are the predictor variables
- The logit function transforms probabilities $(0, 1)$ to the real line $(-\infty, \infty)$

VIF (Variance Inflation Factor) in Logistic Regression

VIF measures multicollinearity among independent variables - how much one predictor can be explained by other predictors in the model.

Formula:
For each independent variable X_j: $\text{VIF}_j = \frac{1}{1 - R_j^2}$

where R_j^2 is the coefficient of determination when X_j is regressed on all other independent variables.

Interpretation:
- **VIF = 1**: No correlation with other predictors (ideal)
- **VIF = 1-5**: Moderate correlation (generally acceptable)
- **VIF = 5-10**: High correlation (concerning - potential multicollinearity problem)
- **VIF > 10**: Severe multicollinearity (problematic - action needed)

NOTE: VIF applies to both linear and logistic regression because it measures relationships among the independent variables, not between predictors and the outcome.

Important question

Output from `lm` command in R:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	(A)	20.673	6.578	3.26e-08 ***
SocialMedia	(B)	4.798	3.172	0.00264 **

Residual standard error: 80.97 on 48 degrees of freedom
Multiple R-squared: (C), Adjusted R-squared: 0.1561
F-statistic: (D) on 1 and 48 DF, p-value: 0.002637

Output from `anova` command in R:

	Df	Sum Sq
SocialMedia	1	65987
Residuals	48	314733

We need to find values of (A), (B), (C), and (D).

$A = b_0 = s.e.(b_0) \times t_{b_0} =$ 135.987	$B = b_1 = s.e.(b_1) \times t_{b_1} =$ 15.21926	$C = R^2 = \frac{SST - SSE}{SST} = \frac{65987}{380720} =$ 0.1733216	$D = F_{\text{obs}} = \frac{MSR}{MSE} = \frac{65987}{6556.938} =$ 10.06369

For C and D: Regression ANOVA table:

Source of variation	Degrees of freedom	Sum of squares	Mean square	F statistic
Regression	1	$SSR =$ 65987	$MSR = \frac{SSR}{1} =$ 65987	$D = F_{\text{obs}} = \frac{MSR}{MSE} = \frac{65987}{6556.938} =$ 10.06369
Residual or Error	$48 = n - 2$	$SSE =$ 314733	$MSE = \frac{SSE}{48} = \frac{314733}{48} =$ 6556.938	
Total	$49 = n - 1$	$SST =$ 380720		

Summary Table

	Small Sample Size	Large Sample Size
Normal + Known Variance	z	z
Normal + Unknown Variance	t	t
Non-Normal + Known Variance	N/A	z
Non-Normal + Unknown Variance	N/A	t

checking for unbiased estimator... An estimator $\hat{\theta}$ is unbiased for parameter θ if: $E\left[\hat{\theta}\right] = \theta$. If $E\left[\hat{Q}\right] \neq Q$, the estimator is biased, with bias given by: $\text{Bias}\left(\hat{Q}\right) = E\left[\hat{Q}\right] - Q$.

MSE measures the average squared difference between the estimator and the true parameter value: $\text{MSE}\left(\hat{Q}\right) = E\left[\left(\hat{Q} - Q\right)^2\right]$. It can be decomposed into variance and bias components: $\text{MSE}\left(\hat{Q}\right) = \text{Var}\left(\hat{Q}\right) + \left(\text{Bias}\left(\hat{Q}\right)\right)^2$. MSE measures the efficiency of an estimator; lower MSE indicates a more efficient estimator.

MME Estimates for Normal Distribution

Given: $X_1, X_2, ..., X_n$ are IID from $N(\mu, \sigma^2)$

Population moments:
- First moment: $E[X] = \mu$
- Second moment: $E[X^2] = \mu^2 + \sigma^2$

Sample moments:
- First sample moment: $\overline{X} = \left(\frac{1}{n}\right) \sum_{i=1}^{n} X_i$
- Second sample moment: $\left(\frac{1}{n}\right) \sum_{i=1}^{n} X_i^2$

MME estimates (equate sample and population moments):

$$\hat{\mu}_{\text{MME}} = \overline{X} = \left(\frac{1}{n}\right) \sum_{i=1}^{n} X_i$$

$$\hat{\sigma}^2_{\text{MME}} = \left(\frac{1}{n}\right) \sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2$$

Properties:
- $\hat{\mu}_{\text{MME}}$ is unbiased: $E[\hat{\mu}] = \mu$ ✓
- $\hat{\sigma}^2_{\text{MME}}$ is BIASED: $E\left[\hat{\sigma}^2_{\text{MME}}\right] = \left(\frac{n-1}{n}\right)\sigma^2 \neq \sigma^2$ with bias = $-\frac{\sigma^2}{n}$
- Unbiased estimator for σ^2 uses $(n-1)$: $S^2 = \left(\frac{1}{n-1}\right) \sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2$
- For Normal distribution: MME = MLE for μ and σ^2

#	What we want to test?	Probability distribution of the statistic	Degrees of freedom	Test Statistic
1.	Population Mean	Normal / t distribution	n-1 for t	$z = \frac{\overline{x}-\mu_0}{\frac{\sigma}{\sqrt{n}}}$ or $t = \frac{\overline{x}-\mu_0}{\frac{s}{\sqrt{n}}}$
2.	Population Proportion	...	...	$z = \frac{\hat{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$ where $\hat{p} = \frac{x}{n}$
3.	Population Variance	Chi-square distribution	n-1	$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$
4.	mean of differences (paired t test)	t distribution	n-1	$t = \frac{\overline{d}-\mu_{d_0}}{\frac{s_d}{\sqrt{n}}}$
5.	test of difference in means of 2 populations (assume equal variance)	Normal / t distribution	$n_1 + n_2 - 2$ for t	$t, z = \frac{\overline{x}_1-\overline{x}_2}{\sqrt{s_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$ (for pooled variance) $t, z = \frac{\overline{x}_1-\overline{x}_2}{\sqrt{\left(\frac{s_1^2}{n_1}\right)+\left(\frac{s_2^2}{n_2}\right)}}$ (for unpooled variance)
6.	test of difference in proportions of 2 populations	...	...	$z_{\text{calc}} = \frac{\widehat{p_1}-\widehat{p_2}}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$
7.	test of difference in variances of 2 populations	F distribution	$n_1 - 1$ and $n_2 - 1$	$F = \frac{s_1^2}{s_2^2}$
8.	difference in means of more than 2 populations (ANOVA)	F distribution	$k - 1$ and $N - k$	$F = \frac{\text{MS}_{\text{between}}}{\text{MS}_{\text{within}}}$
9.	barletts test for equal variances	Chi-square distribution	$k - 1$	$\chi^2 = \frac{\left(\ln(S_p^2)-\sum\left((n_i-1)\frac{\ln\left(s_i^2\right)}{n-k}\right)\right)*(n-k)}{1+\left(\frac{1}{3(k-1)}\right)*\left(\sum\left(\frac{1}{n_i-1}\right)-\left(\frac{1}{n-k}\right)\right)}$
10.	two way ANOVA	F distribution	$(r-1), (c-1), (\text{rc}(n-1))$	$F = \frac{\text{MS}_{\text{factor}}}{\text{MS}_{\text{error}}}$
11.	goodness of fit	Chi-square distribution	$c - 1$	$\chi^2 = \sum\left(\frac{(O_i-E_i)^2}{E_i}\right)$
12.	test of independence	Chi-square distribution	(r - 1)(c - 1)	$\chi^2 = \sum\left(\frac{(O_i-E_i)^2}{E_i}\right)$
13.	wilcoxon signed rank test	approximate Normal distribution for large n	...	$Z = \frac{W-\left(\frac{n(n+1)}{4}\right)}{\sqrt{n(n+1)\frac{2n+1}{24}}}$
14.	wilcoxon mann whitney test	approximate Normal distribution for large n	...	$Z = \frac{U-\left(n_1\frac{n_2}{2}\right)}{\sqrt{n_1 n_2 \frac{n_1+n_2+1}{12}}}$
15.	kruskal wallies test	approximate χ^2 distribution for large n $H \sim \chi^2_{\text{k-1}}$	$k - 1$	$H = \left(\frac{12}{N(N+1)}\right)\sum\left(\left(\frac{R_j^2}{n_j}\right)\right) - 3(N+1)$

NOTE: For χ^2 tests, ensure that the expected frequencies are sufficiently large (usually at least 5) to validate the use of the chi-square approximation.