**Common Symbols**

n = Sample Size — N = Population Size

$\bar{x}$ = Sample Mean — $\mu$ = Population Mean

s = Sample Standard Deviation — $\sigma$ = Population Standard Deviation

$s^2$ = Sample Variance — $\sigma^2$ = Population Variance

$\hat{p}$ = Sample Proportion — p = Population Proportion

r = Sample Correlation Coefficient — $\rho$ = Population Correlation Coefficient

**Descriptive Statistics** — **In Excel use Data > Data Analysis > Descriptive Statisitcs**

$$\bar{x} = \frac{\Sigma x}{n} \qquad s^2 = \frac{\Sigma(x-\bar{x})^2}{n-1} \qquad s = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}} \qquad z = \frac{x-\bar{x}}{s}$$

Range = Max – Min

IQR=interquartile range = $Q_3 - Q_1$

Coefficient of Variation = $\frac{s}{\bar{x}}(100\%)$

Finding Outlier Limits: Lower = $Q_1$ – (1.5*IQR) Upper = $Q_3$ + (1.5*IQR)

**Probability**

Union Rule: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Complement Rule: $P(A^C) = 1 - P(A)$

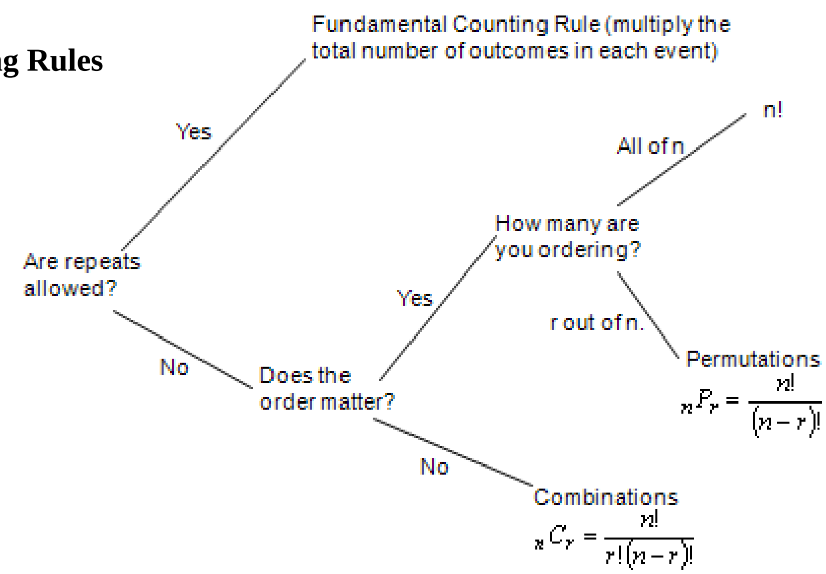
Conditional Probability: $P(A \mid B) = \frac{P(A \cap B)}{P(B)}$

Mutually Exclusive Events: $P(A \cap B) = 0$

Dependent Events: $P(A \cap B) = P(B) \cdot P(A \mid B)$

Independent Events: $P(A \cap B) = P(B) \cdot P(A)$

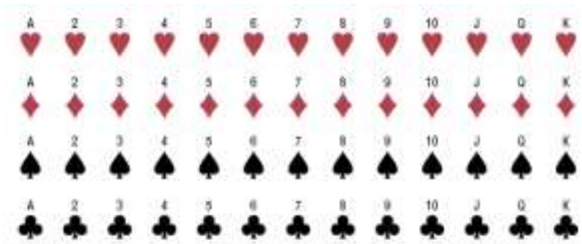
**Counting Rules**




**Sum of Two Dice**

| | | Second Die | | | | | |
|---|---|---|---|---|---|---|---|
| | + | 1 | 2 | 3 | 4 | 5 | 6 |
| First Die | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| | 6 | 7 | 8 | 9 | 10 | 11 | 12 |

**Standard Deck of Cards**




Face Cards are Jack (J), Queen (Q) and King (K)

## Discrete Probability Distributions

$\mathrm{E}(x) = \mu = \Sigma x \cdot P(x) \qquad \sigma^2 = \left(\Sigma x^2 \cdot P(x)\right) - \mu^2 \qquad \sigma = \sqrt{\left(\Sigma x^2 \cdot P(x)\right) - \mu^2}$

**Binomial Distribution**: $P(x) = {}_n\mathrm{C}_x \cdot p^x \cdot q^{(n-x)} \qquad \mu = n \cdot p \qquad \sigma^2 = n \cdot p \cdot q \qquad \sigma = \sqrt{n \cdot p \cdot q}$

| **Signs Are Important For Discrete Distributions!** | | |
|---|---|---|
| **P(X=x)** | **P(X≤x)** | **P(X≥x)** |
| Is | Is less than or equal to | Is greater than or equal to |
| Is equal to | Is at most | Is at least |
| Is exactly the same as | Is not greater than | Is not less than |
| Has not changed from | Within | |
| Is the same as | | |
| | | |
| =BINOM.DIST(x,n,p,false) | =BINOM.DIST(x,n,p,true) | =1-BINOM.DIST(x–1,n,p,true) |
| TI: binompdf(n,p,x) | TI: binomcdf(n,p,x) | TI: 1–binomcdf(n,p,x–1) |
| | | |
| **Where:** | **P(X>x)** | **P(X<x)** |
| x is the value in the | More than | Less than |
| question that you are | Greater than | Below |
| finding the probability for. | Above | Lower than |
| p is the proportion of a | Higher than | Shorter than |
| success expressed as a | Longer than | Smaller than |
| decimal between 0 and 1 | Bigger than | Decreased |
| n is the sample size | Increased | Reduced |
| | Smaller | Larger |
| | | |
| Excel: | =1–BINOM.DIST(x,n,p,true) | =BINOM.DIST(x–1,n,p,true) |
| TI-Calculator: | 1–binomcdf(n,p,x) | binomcdf(n,p,x–1) |

**Poisson Distribution**: $P(x) = \frac{e^{-\mu}\mu^x}{x!}$ Change mean to fit the units in the question: $\text{new } \mu = \text{old } \mu \times \left(\frac{\text{new units}}{\text{old units}}\right)$

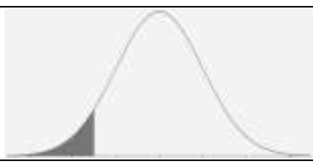
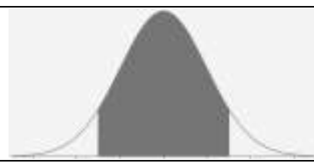
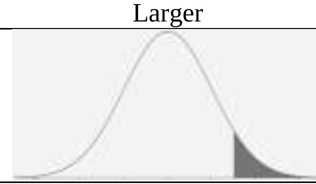
| **P(X=x)** | **P(X≤x)** | **P(X≥x)** |
|---|---|---|
| Is | Is less than or equal to | Is greater than or equal to |
| Is equal to | Is at most | Is at least |
| Is exactly the same as | Is not greater than | Is not less than |
| Has not changed from | Within | |
| Is the same as | | |
| | | |
| =POISSON.DIST(x,mean,false) | =POISSON.DIST(x,mean,true) | =1–POISSON.DIST(x–1,mean,true) |
| TI: poissonpdf(mean,x) | TI: poissoncdf(mean,x) | TI: 1–poissoncdf(mean,x–1) |
| | | |
| **Where:** | **P(X>x)** | **P(X<x)** |
| x is the value in the | More than | Less than |
| question that you are | Greater than | Below |
| finding the probability for. | Above | Lower than |
| | Higher than | Shorter than |
| The mean has been rescaled to | Longer than | Smaller than |
| the units of the question. | Bigger than | Decreased |
| | Increased | Reduced |
| | Smaller | Larger |
| | | |
| Excel: | =1–POISSON.DIST(x,mean,true) | =POISSON.DIST(x–1,mean,true) |
| TI-Calculator: | 1–poissoncdf(mean,x) | poissoncdf(mean,x–1) |

## Continuous Distributions

$z = \frac{x-\mu}{\sigma}$ $\quad x = z \cdot \sigma + \mu$ $\quad$ Central Limit Theorem: $z = \frac{\bar{x}-\mu}{\left(\frac{\sigma}{\sqrt{n}}\right)}$

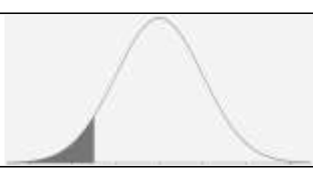
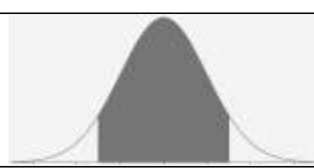
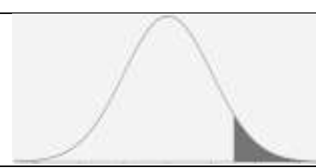
Note that for a continuous distribution there is no area at a line under the curve, so ≥ and > will have the same probability and use the same Excel commands.

## Normal Distribution

| Normal Distribution Finding a Probability | | |
|---|---|---|
| **$P(X \le x)$ or $P(X<x)$** | **$P(x_1<X<x_2)$ or $P(x_1 \le X \le x_2)$** | **$P(X \ge x)$ or $P(X>x)$** |
| Is less than or equal to | Between | Is greater than or equal to |
| Is at most | | Is at least |
| Is not greater than | | Is not less than |
| Within | | More than |
| Less than | | Greater than |
| Below | | Above |
| Lower than | | Higher than |
| Shorter than | | Longer than |
| Smaller than | | Bigger than |
| Decreased | | Increased |
| Reduced | | Larger |
|  |  |  |
| =NORM.DIST (x,μ ,σ,true) | =NORM.DIST ($x_2$,μ ,σ,true)- NORM.DIST ($x_1$,μ ,σ,true) | =1- NORM.DIST (x,μ ,σ,true) |

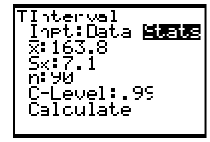
Note that the NORM.S.DIST function is for a standard normal when μ=0 and σ=1.

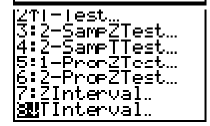
## Inverse Normal Distribution

| Normal Distribution Finding an X-value Given an Area or Probability | | |
|---|---|---|
| **$P(X \le x)$ or $P(X<x)$** | **$P(x_1<X<x_2)$ or $P(x_1 \le X \le x_2)$** | **$P(X \ge x)$ or $P(X>x)$** |
| Lower | Between | Upper |
| Bottom | | Top |
| Below | | Above |
| Reduced | | More than |
| Less than | | Greater than |
| Lower than | | Larger |
| Shorter than | | Higher than |
| Smaller than | | Longer than |
| Decreased | | Bigger than |
| | | Increased |
| | | |
|  |  |  |
| =NORM.INV(area,μ,σ) | $x_1$ =NORM.INV(1-area/2,μ ,σ) $x_2$ = \| $x_1$\| | =NORM.INV(1-area, μ,σ) |

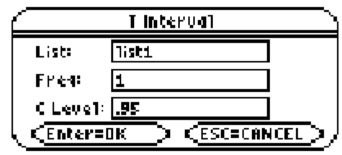
**Confidence Intervals** **When n < 30 the variable must be approximately normally distributed.**

**t-Interval** **The 100(1 - α)% confidence interval for μ, σ is unknown, is** $\bar{x} \pm t_{n-1,\alpha/2}\frac{s}{\sqrt{n}}$ **.**

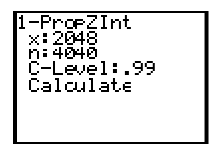


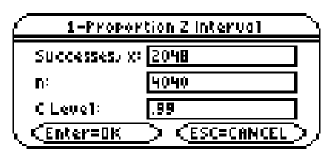

- On the TI-83 you can find a confidence interval using the statistics menu. Press the [STAT] key, arrow over to the [TESTS] menu, arrow down to the [8:TInterval] option and press the [ENTER] key. Arrow over to the [Stats] menu and press the [ENTER] key. Then type in the mean, sample standard deviation, sample size and confidence level, arrow down to [Calculate] and press the [ENTER] key. The calculator returns the answer in interval notation. Be careful, if you accidentally use the [7:ZInterval] option you would get the wrong answer.
- *Or* (If you have raw data in list one) Arrow over to the [Data] menu and press the [ENTER] key. Then type in the list name, $L_1$, leave Freq:1 alone, enter the confidence level, arrow down to [Calculate] and press the [ENTER] key.
- On the TI-89 go to the [Apps] **Stat/List Editor**, then select 2nd then F7 [Ints], then select **1:TInterval**. Choose the input method, data is when you have entered **data** into a list previously or **stats** when you are given the mean and standard deviation already. Type in the mean, standard deviation, sample size (or list name (list1), and Freq: 1) and confidence level, and press the [ENTER] key. The calculator returns the answer in interval notation. Be careful, if you accidentally use the [1:ZInterval] option you would get the wrong answer.




**1 Proportion z-Interval** $\hat{p} \pm z_{\alpha/2}\sqrt{\left(\frac{\hat{p}(1-\hat{p})}{n}\right)}$



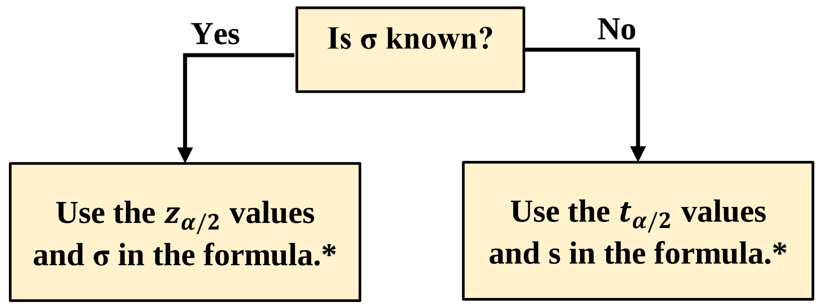

- On the TI-84 press the [STAT] key, arrow over to the [TESTS] menu, arrow down to the [A:1-PropZInterval] option and press the [ENTER] key. Then type in the values for X, sample size and confidence level, arrow down to [Calculate] and press the [ENTER] key. The calculator returns the answer in interval notation. Note: sometimes you are not given the x value but a percentage instead. To find the x to use in the calculator, multiply $\hat{p}$ by the sample size and round off to the nearest integer. The calculator will give you an error message if you put in a decimal for x or n. For example if $\hat{p}$ = .22 and n = 124 then .22*124 = 27.28, so use x = 27.
- On the TI-89 go to the [Apps**] Stat/List Editor**, then select 2nd then F7 [Ints], then select **5: 1-PropZInt**. Type in the values for X, sample size and confidence level, and press the [ENTER] key. The calculator returns the answer in interval notation. Note: sometimes you are not given the x value but a percentage instead. To find the x value to use in the calculator, multiply $\overline{p}$ by the sample size and round off to the nearest integer. The calculator will give you an error message if you put in a decimal for x or n. For example if $\overline{p}$ = .22 and n = 124 then .22*124 = 27.28, so use x = 27.




**Sample Size** $n = \left(\frac{z_{\alpha/2}\sigma}{E}\right)^2$ $n = \hat{p}(1-\hat{p})\left(\frac{z_{\alpha/2}}{E}\right)^2$ Always round n up to the next integer.

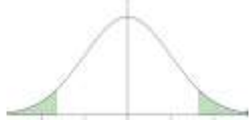
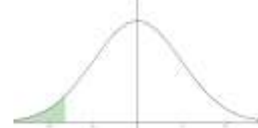
**Hypothesis Testing**

α = P(Type I error)



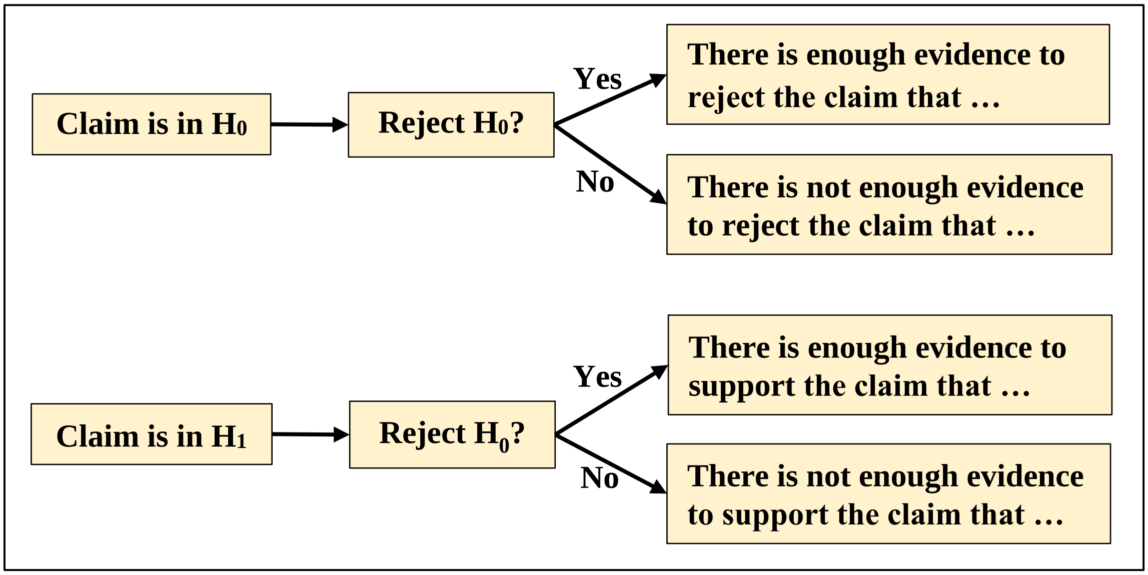

***If n < 30, the variable must be normally**

**Look for these key words to help set up your hypotheses:**

| Two-tailed Test | Right-tailed Test | Left-tailed Test |
|---|---|---|
| $H_0: \mu = \mu_0 \quad H_1: \mu \neq \mu_0$ | $H_0: \mu = \mu_0 \quad H_1: \mu > \mu_0$ | $H_0: \mu = \mu_0 \quad H_1: \mu < \mu_0$ |
|  |  |  |
| **Claim is in the Null Hypothesis** | | |
| = | ≤ | ≥ |
| Is equal to | Is less than or equal to | Is greater than or equal to |
| Is exactly the same as | Is at most | Is at least |
| Has not changed from | Is not more than | Is not less than |
| Is the same as | Within | |
| **Claim is in the Alternative Hypothesis** | | |
| ≠ | > | < |
| Is not | More than | Less than |
| Is not equal to | Greater than | Below |
| Is different from | Above | Lower than |
| Has changed from | Higher than | Shorter than |
| Is not the same as | Longer than | Smaller than |
| | Bigger than | Decreased |
| | Increased | Reduced |

**The rejection rule:**

- p-value method: reject $H_0$ when the p-value ≤ α.
- Critical value method: reject $H_0$ when the test statistic is in the critical tail(s).
- Confidence Interval method, reject $H_0$ when the hypothesized value found in $H_0$ is outside the bounds of the confidence interval.




Finish conclusion with context and units from question.

**One Sample Tests:**

**1-Sample Mean t-test:** $H_0: \mu = \mu_0$, $H_1: \mu \neq \mu_0$ Test statistic when σ is unknown: $t = \dfrac{\bar{x} - \mu_0}{\left(s/\sqrt{n}\right)}$ with df= n – 1

**1-Sample Proportion z-test:** $H_0: p = p_0$, $H_1: p \neq p_0$ Test statistic is $z = \dfrac{\hat{p} - p}{\sqrt{\dfrac{p \cdot q}{n}}}$ where $\hat{p} = \dfrac{x}{n}$

**Two Sample Tests:**

**2-Sample Means t-test – Independent Populations**

$H_0: \mu_1 = \mu_2$, $H_1: \mu_1 \neq \mu_2$ or $H_0: \mu_1 - \mu_2 = (\mu_1 - \mu_2)_0$, $H_1: \mu_1 - \mu_2 \neq (\mu_1 - \mu_2)_0$ where usually you have $H_0: \mu_1 - \mu_2 = 0$, $H_1: \mu_1 - \mu_2 \neq 0$

Note that $(\mu_1 - \mu_2)_0 = 0$ in most cases

**Assuming Equal Variances**

Test statistic when $\sigma_1^2$ and $\sigma_2^2$ are unknown: $t = \dfrac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)_0}{\sqrt{\left(\left(\dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}\right)\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)\right)}}$

**Assuming Unequal Variances**

Test statistic when $\sigma_1^2$ and $\sigma_2^2$ are unknown: $t = \dfrac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)_0}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$ with

$$df = \frac{\left(\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}\right)^2}{\left(\left(\dfrac{1}{n_1 - 1}\left(\dfrac{s_1^2}{n_1}\right)^2\right) + \left(\dfrac{1}{n_2 - 1}\left(\dfrac{s_2^2}{n_2}\right)^2\right)\right)}$$

**2-Sample Means t-test – Dependent Populations**

Find the difference (d) between each matched pairs. $H_0: \mu_D = 0$, $H_1: \mu_D \neq 0$ Test statistic: $t = \dfrac{\bar{d} - \mu_0}{\left(s_d/\sqrt{n}\right)}$

**2 Proportions**

$H_0: p_1 = p_2$

$H_1: p_1 \neq p_2$

Test statistic $z = \dfrac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)_0}{\sqrt{\hat{p}\hat{q}\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}$, (usually $(p_1 - p_2)_0 = 0$) where $\hat{p} = \dfrac{x_1 + x_2}{n_1 + n_2}$, $\hat{q} = 1 - \hat{p}$

**Correlation and Regression**

$H_0: \rho = 0$

$H_1: \rho \neq 0$

Test statistic for correlation: $t = r\sqrt{\left(\dfrac{(n-2)}{(1-r^2)}\right)}$ with df= n – 2 $\hat{y} = a + bx$

r = sample correlation coefficient $\rho$ = population correlation coefficient
a = y-intercept b = slope s = $s_{est}$ = standard error of estimate $R^2$ = coefficient of determination

**One-Factor ANOVA** table k=#of groups, N=total of all n's

$H_0$: $\mu_1=\mu_2=\mu_3=\ldots=\mu_k$

$H_1$: At least one mean is different CV: Always a right-tailed F, use Excel =F.INV.RT(α, $df_B$, $df_W$)

ANOVA Table

| Source | SS | df | MS | F (Test Statistic) |
|---|---|---|---|---|
| Between (Treatment or Factor) | $SS_B=\Sigma n(\bar{x} - \bar{x}_{GM})^2$ | k–1 | $MS_B = SS_B / df_B$ | $F = MS_B / MS_W$ |
| Within (Error) | $SS_W=\Sigma(n-1)s^2$ | N–k | $MS_W = SS_W / df_W$ | |
| Total | $SS_T$ | N–1 | | |

When you reject $H_0$ for a one-factor ANOVA then you should do a multiple comparison. For example for 3 groups you would have the following 3 comparisons. (4 groups would have ${}_4C_2$=6 comparisons)

$H_0:\mu_1=\mu_2$ $H_0:\mu_1=\mu_3$ $H_0:\mu_2=\mu_3$

$H_1:\mu_1\neq\mu_2$ $H_1:\mu_1\neq\mu_3$ $H_1:\mu_2\neq\mu_3$

- Bonferroni Test $t = \dfrac{(\bar{x}_i - \bar{x}_j)}{\sqrt{\left(MSE\left(\dfrac{1}{n_i} + \dfrac{1}{n_j}\right)\right)}}$ with df = N – k and to get the p-value you would multiply the tail areas by ${}_kC_2$ groups. For example if you use the tcdf in your calculator to find the area in both the tails and you have 4 groups you would multiply the tail areas by ${}_4C_2$=6.