Useful Formulas

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \qquad s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2} \qquad r = \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{x_i - \bar{x}}{s_x}\right)\left(\frac{y_i - \bar{y}}{s_y}\right)$$

$$\hat{y} = b_0 + b_1 x \qquad b_1 = \frac{r s_y}{s_x} \qquad b_0 = \bar{y} - b_1\bar{x}$$

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B) \qquad P(A \text{ and } B) = P(A)P(B|A)$$

$$P(B|A) = \frac{P(A \text{ and } B)}{P(A)} \qquad P(A) = P(B)P(A|B) + P(B^c)P(A|B^c)$$

$$P(B|A) = \frac{P(B)P(A|B)}{P(B)P(A|B) + P(B^c)P(A|B^c)}$$

	Discrete case	**Continuous case**
$P(a \le X \le b)$	$\sum_{i: a \le x_i \le b} P(X = x_i)$	$\int_a^b f(x)\,dx$
$\mu = E[X]$	$\sum_i x_i P(X = x_i)$	$\int_{-\infty}^{\infty} x f(x)\,dx$
$\text{Var}(X)$	$\sum_i (x_i - \mu)^2 P(X = x_i)$	$\left(\int_{-\infty}^{\infty} x^2 f(x)\,dx\right) - \mu^2$
$\text{SD}(X)$	$\sqrt{\text{Var}(X)}$	$\sqrt{\text{Var}(X)}$

Distribution	**probability function or density**	$E[X]$	$\text{SD}(X)$
Binomial(n, p)	$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$ for $k = 0, 1, ..., n$	np	$\sqrt{np(1-p)}$
Poisson(λ)	$P(X = k) = e^{-\lambda}\lambda^k/k!$ for $k = 0, 1, \ldots$	λ	$\sqrt{\lambda}$
Geometric(p)	$P(X = k) = (1-p)^{k-1}p$ for $k = 1, 2, ...$	$1/p$	$\sqrt{1-p}/p$
Uniform(a, b)	$f(x) = 1/(b-a)$ for $a < x < b$	$(a+b)/2$	$(b-a)/\sqrt{12}$
Exponential(λ)	$f(x) = \lambda e^{-\lambda x}$ for $x > 0$	$1/\lambda$	$1/\lambda$
Normal(μ, σ^2)	$f(x) = (2\pi\sigma^2)^{-1/2} e^{-(x-\mu)^2/2\sigma^2}$	μ	σ

$$E[X + c] = E[X] + c \qquad E[cX] = cE[X] \qquad E[X + Y] = E[X] + E[Y].$$

$$\text{Var}(X + c) = \text{Var}(X) \qquad \text{Var}(cX) = c^2\text{Var}(X)$$

$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$ if X and Y are independent.

$$E[\bar{X}] = \mu \qquad \text{SD}(\bar{X}) = \sigma/\sqrt{n} \qquad E[\hat{p}] = p \qquad \text{SD}(\hat{p}) = \sqrt{\frac{p(1-p)}{n}}$$

Test	**ME for Interval**	**Test Statistic**	**df**
One-proportion z-test	$z^*\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$	$Z = \frac{p - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$	N/A
One-sample t-test	$t^*\frac{s}{\sqrt{n}}$	$T = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$	$n-1$
Paired t-test	$t^*\frac{s_d}{\sqrt{n}}$	$T = \frac{\bar{d}}{s_d/\sqrt{n}}$	$n-1$
Two-sample t-test	$t^*\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$	$T = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$	$\min\{n_1 - 1, n_2 - 1\}$
Regression slope	$t^* SE(b_1)$	$T = \frac{b_1}{SE(b_1)}$	$n-2$
Chi-square test	N/A	$\chi^2 = \sum \frac{(\text{Obs - Exp})^2}{\text{Exp}}$	see below

Chi-square test for goodness-of-fit: df = number of categories - 1

Chi-square test for independence/homogeneity: df = (rows - 1)(columns - 1)

Expected counts in chi-square test for independence/homogeneity: $\frac{\text{Row Total} \times \text{Column Total}}{\text{Table Total}}$

$$SE(\hat{\mu}_y) = \sqrt{\frac{s^2}{n} + (x^* - \bar{x})^2 SE(b_1)^2} \qquad SE(\hat{y}) = \sqrt{s^2 + \frac{s^2}{n} + (x^* - \bar{x})^2 SE(b_1)^2}$$

$$SE(b_1) = \frac{s}{s_x\sqrt{n-1}}$$