Formulas and Tables by Mario F. Triola

Ch. 3: Descriptive Statistics

$\bar{x} = \frac{\Sigma x}{n}$ Mean

$\bar{x} = \frac{\Sigma(f \cdot x)}{\Sigma f}$ Mean (frequency table)

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

$s = \sqrt{\frac{n(\Sigma x^2) - (\Sigma x)^2}{n(n - 1)}}$ Standard deviation (shortcut)

$s = \sqrt{\frac{n[\Sigma(f \cdot x^2)] - [\Sigma(f \cdot x)]^2}{n(n - 1)}}$ Standard deviation (frequency table)

variance $= s^2$

Ch. 4: Probability

$P(A \text{ or } B) = P(A) + P(B)$ if A, B are mutually exclusive

$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$
if A, B are not mutually exclusive

$P(A \text{ and } B) = P(A) \cdot P(B)$ if A, B are independent

$P(A \text{ and } B) = P(A) \cdot P(B|A)$ if A, B are dependent

$P(\bar{A}) = 1 - P(A)$ Rule of complements

${}_nP_r = \frac{n!}{(n - r)!}$ Permutations (no elements alike)

$\frac{n!}{n_1!\, n_2! \ldots n_k!}$ Permutations (n_1 alike, . . .)

${}_nC_r = \frac{n!}{(n - r)!\, r!}$ Combinations

Ch. 5: Probability Distributions

$\mu = \Sigma[x \cdot P(x)]$ Mean (prob. dist.)

$\sigma = \sqrt{\Sigma[x^2 \cdot P(x)] - \mu^2}$ Standard deviation (prob. dist.)

$P(x) = \frac{n!}{(n - x)!\, x!} \cdot p^x \cdot q^{n-x}$ Binomial probability

$\mu = n \cdot p$ Mean (binomial)

$\sigma^2 = n \cdot p \cdot q$ Variance (binomial)

$\sigma = \sqrt{n \cdot p \cdot q}$ Standard deviation (binomial)

$P(x) = \frac{\mu^x \cdot e^{-\mu}}{x!}$ Poisson distribution where $e = 2.71828$

Ch. 6: Normal Distribution

$z = \frac{x - \mu}{\sigma}$ or $\frac{x - \bar{x}}{s}$ Standard score

$\mu_{\bar{x}} = \mu$ Central limit theorem

$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ Central limit theorem (Standard error)

Ch. 7: Confidence Intervals (one population)

$\hat{p} - E < p < \hat{p} + E$ Proportion

where $E = z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}$

$\bar{x} - E < \mu < \bar{x} + E$ Mean

where $E = t_{\alpha/2}\frac{s}{\sqrt{n}}$ (σ unknown)

or $E = z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ (σ known)

$\frac{(n - 1)s^2}{\chi^2_R} < \sigma^2 < \frac{(n - 1)s^2}{\chi^2_L}$ Variance

Ch. 7: Sample Size Determination

$n = \frac{[z_{\alpha/2}]^2 0.25}{E^2}$ Proportion

$n = \frac{[z_{\alpha/2}]^2 \hat{p}\hat{q}}{E^2}$ Proportion ($\hat{p}$ and $\hat{q}$ are known)

$n = \left[\frac{z_{\alpha/2}\sigma}{E}\right]^2$ Mean

Ch. 8: Test Statistics (one population)

$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$ Proportion—one population

$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$ Mean—one population (σ unknown)

$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$ Mean—one population (σ known)

$\chi^2 = \frac{(n - 1)s^2}{\sigma^2}$ Standard deviation or variance—one population

Formulas and Tables by Mario F. Triola

Ch. 9: Confidence Intervals (two populations)

$(\hat{p}_1 - \hat{p}_2) - E < (p_1 - p_2) < (\hat{p}_1 - \hat{p}_2) + E$

where $E = z_{\alpha/2}\sqrt{\dfrac{\hat{p}_1\hat{q}_1}{n_1} + \dfrac{\hat{p}_2\hat{q}_2}{n_2}}$

$(\bar{x}_1 - \bar{x}_2) - E < (\mu_1 - \mu_2) < (\bar{x}_1 - \bar{x}_2) + E$ (Indep.)

where $E = t_{\alpha/2}\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}$ (df = smaller of $n_1 - 1, n_2 - 1$)

(σ_1 and σ_2 unknown and not assumed equal)

$E = t_{\alpha/2}\sqrt{\dfrac{s_p^2}{n_1} + \dfrac{s_p^2}{n_2}}$ (df $= n_1 + n_2 - 2$)

$s_p^2 = \dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 - 1) + (n_2 - 1)}$

(σ_1 and σ_2 unknown but assumed equal)

$E = z_{\alpha/2}\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}$

(σ_1, σ_2 known)

$\bar{d} - E < \mu_d < \bar{d} + E$ (Matched pairs)

where $E = t_{\alpha/2}\dfrac{s_d}{\sqrt{n}}$ (df $= n - 1$)

Ch. 9: Test Statistics (two populations)

$z = \dfrac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\dfrac{\bar{p}\,\bar{q}}{n_1} + \dfrac{\bar{p}\,\bar{q}}{n_2}}}$ Two proportions, $\bar{p} = \dfrac{x_1 + x_2}{n_1 + n_2}$

$t = \dfrac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$ df = smaller of $n_1 - 1, n_2 - 1$

Two means—independent; σ_1 and σ_2 unknown, and not assumed equal.

$t = \dfrac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_p^2}{n_1} + \dfrac{s_p^2}{n_2}}}$ (df $= n_1 + n_2 - 2$)

$s_p^2 = \dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$

Two means—independent; σ_1 and σ_2 unknown, but assumed equal.

$z = \dfrac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$ Two means—independent; σ_1, σ_2 known.

$t = \dfrac{\bar{d} - \mu_d}{\dfrac{s_d}{\sqrt{n}}}$ Two means—matched pairs (df $= n - 1$)

$F = \dfrac{s_1^2}{s_2^2}$ Standard deviation or variance—two populations (where $s_1^2 \geq s_2^2$)

Ch. 10: Linear Correlation/Regression

Correlation $r = \dfrac{n\Sigma xy - (\Sigma x)(\Sigma y)}{\sqrt{n(\Sigma x^2) - (\Sigma x)^2}\sqrt{n(\Sigma y^2) - (\Sigma y)^2}}$

or $r = \dfrac{\Sigma(z_x z_y)}{n - 1}$ where z_x = z score for x, z_y = z score for y

Slope: $b_1 = \dfrac{n\Sigma xy - (\Sigma x)(\Sigma y)}{n(\Sigma x^2) - (\Sigma x)^2}$ or $b_1 = r\dfrac{s_y}{s_x}$

y-Intercept:

$b_0 = \bar{y} - b_1\bar{x}$ or $b_0 = \dfrac{(\Sigma y)(\Sigma x^2) - (\Sigma x)(\Sigma xy)}{n(\Sigma x^2) - (\Sigma x)^2}$

$\hat{y} = b_0 + b_1 x$ Estimated eq. of regression line

$r^2 = \dfrac{\text{explained variation}}{\text{total variation}}$

$s_e = \sqrt{\dfrac{\Sigma(y - \hat{y})^2}{n - 2}}$ or $\sqrt{\dfrac{\Sigma y^2 - b_0\Sigma y - b_1\Sigma xy}{n - 2}}$

$\hat{y} - E < y < \hat{y} + E$ Prediction interval

where $E = t_{\alpha/2}s_e\sqrt{1 + \dfrac{1}{n} + \dfrac{n(x_0 - \bar{x})^2}{n(\Sigma x^2) - (\Sigma x)^2}}$

Ch. 11: Goodness-of-Fit and Contingency Tables

$\chi^2 = \Sigma\dfrac{(O - E)^2}{E}$ Goodness-of-fit (df $= k - 1$)

$\chi^2 = \Sigma\dfrac{(O - E)^2}{E}$ Contingency table [df $= (r - 1)(c - 1)$]

where $E = \dfrac{(\text{row total})(\text{column total})}{(\text{grand total})}$

$\chi^2 = \dfrac{(|b - c| - 1)^2}{b + c}$ McNemar's test for matched pairs (df $= 1$)

Ch. 12: One-Way Analysis of Variance

Procedure for testing H_0: $\mu_1 = \mu_2 = \mu_3 = \ldots$

1. Use software or calculator to obtain results.
2. Identify the P-value.
3. Form conclusion:

   If P-value $\leq \alpha$, reject the null hypothesis of equal means.
   If P-value $> \alpha$, fail to reject the null hypothesis of equal means.

Ch. 12: Two-Way Analysis of Variance

Procedure:

1. Use software or a calculator to obtain results.
2. Test H_0: There is no interaction between the row factor and column factor.
3. Stop if H_0 from Step 2 is rejected.

   If H_0 from Step 2 is not rejected (so there does not appear to be an interaction effect), proceed with these two tests:
   Test for effects from the row factor.
   Test for effects from the column factor.

Ch. 13: Nonparametric Tests

$$z = \frac{(x + 0.5) - (n/2)}{\frac{\sqrt{n}}{2}}$$ Sign test for $n > 25$

$$z = \frac{T - n(n+1)/4}{\sqrt{\frac{n(n+1)(2n+1)}{24}}}$$ Wilcoxon signed ranks (matched pairs and $n > 30$)

$$z = \frac{R - \mu_R}{\sigma_R} = \frac{R - \frac{n_1(n_1 + n_2 + 1)}{2}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}}$$ Wilcoxon rank-sum (two independent samples)

$$H = \frac{12}{N(N+1)}\left(\frac{R_1^2}{n_1} + \frac{R_2^2}{n_2} + \cdots + \frac{R_k^2}{n_k}\right) - 3(N+1)$$

Kruskal-Wallis (chi-square df $= k - 1$)

$$r_s = 1 - \frac{6\Sigma d^2}{n(n^2 - 1)}$$ Rank correlation

$\left(\text{critical values for } n > 30: \frac{\pm z}{\sqrt{n-1}}\right)$

$$z = \frac{G - \mu_G}{\sigma_G} = \frac{G - \left(\frac{2n_1 n_2}{n_1 + n_2} + 1\right)}{\sqrt{\frac{(2n_1 n_2)(2n_1 n_2 - n_1 - n_2)}{(n_1 + n_2)^2 (n_1 + n_2 - 1)}}}$$ Runs test for $n > 20$

Ch. 14: Control Charts

R chart: Plot sample ranges

UCL: $D_4\overline{R}$

Centerline: $\overline{R}$

LCL: $D_3\overline{R}$

$\overline{x}$ chart: Plot sample means

UCL: $\overline{\overline{x}} + A_2\overline{R}$

Centerline: $\overline{\overline{x}}$

LCL: $\overline{\overline{x}} - A_2\overline{R}$

p chart: Plot sample proportions

UCL: $\overline{p} + 3\sqrt{\frac{\overline{p}\,\overline{q}}{n}}$

Centerline: $\overline{p}$

LCL: $\overline{p} - 3\sqrt{\frac{\overline{p}\,\overline{q}}{n}}$

TABLE A-6 Critical Values of the Pearson Correlation Coefficient r

n	$\alpha = .05$	$\alpha = .01$
4	.950	.990
5	.878	.959
6	.811	.917
7	.754	.875
8	.707	.834
9	.666	.798
10	.632	.765
11	.602	.735
12	.576	.708
13	.553	.684
14	.532	.661
15	.514	.641
16	.497	.623
17	.482	.606
18	.468	.590
19	.456	.575
20	.444	.561
25	.396	.505
30	.361	.463
35	.335	.430
40	.312	.402
45	.294	.378
50	.279	.361
60	.254	.330
70	.236	.305
80	.220	.286
90	.207	.269
100	.196	.256

NOTE: To test H_0: $\rho = 0$ (no correlation) against H_1: $\rho \neq 0$ (correlation), reject H_0 if the absolute value of r is greater than or equal to the critical value in the table.

Control Chart Constants

Subgroup Size n	D_3	D_4	A_2
2	0.000	3.267	1.880
3	0.000	2.574	1.023
4	0.000	2.282	0.729
5	0.000	2.114	0.577
6	0.000	2.004	0.483
7	0.076	1.924	0.419

Inferences about μ: choosing between t and normal distributions		
t distribution:		σ not known and normally distributed population
	or	σ not known and $n > 30$
Normal distribution:		σ known and normally distributed population
	or	σ known and $n > 30$
Nonparametric method or bootstrapping: Population not normally distributed and $n \le 30$		

Procedure for Hypothesis Tests

1. Identify the Claim
Identify the claim to be tested and express it in symbolic form.

↓

2. Give Symbolic Form
Give the symbolic form that must be true when the original claim is false.

↓

3. Identify Null and Alternative Hypothesis
Consider the two symbolic expressions obtained so far:
- **Alternative hypothesis H_1** is the one *NOT* containing equality, so H_1 uses the symbol $>$ or $<$ or $\neq$.
- **Null hypothesis H_0** is the symbolic expression that the parameter equals the fixed value being considered.

↓

4. Select Significance Level
Select the **significance level α** based on the seriousness of a type I error. Make α small if the consequences of rejecting a true H_0 are severe.
- The values of 0.05 and 0.01 are very common.

↓

5. Identify the Test Statistic
Identify the test statistic that is relevant to the test and determine its sampling distribution (such as normal, *t*, chi-square).

P-Value Method

6. Find Values
Find the value of the **test statistic** and the ***P*-value** (see Figure 8-3). Draw a graph and show the test statistic and *P*-value.

↓

7. Make a Decision
- **Reject H_0** if *P*-value $\leq \alpha$.
- **Fail to reject H_0** if *P*-value $> \alpha$.

Critical Value Method

6. Find Values
Find the value of the **test statistic** and the **critical values**. Draw a graph showing the test statistic, critical value(s) and critical region.

↓

7. Make a Decision
- **Reject H_0** if the test statistic is in the critical region.
- **Fail to reject H_0** if the test statistic is not in the critical region.

↓

8. Restate Decision in Nontechnical Terms
Restate this previous decision in simple nontechnical terms, and address the original claim.

Finding *P*-Values