1 Simple Linear Correlation

1.1 Kinds of Correlation

- simple
- multiple
- partial

1.2 Correlation Coefficient

$$r = \frac{\sum xy}{\sqrt{\sum x^2 \sum y^2}}$$

$$-1 \le r \le 1$$

c.f. regression coefficient (b): $-\infty \le b \le \infty$

- b measures effect of X on Y
- r measures strength of association betw X & Y

1.3 Coefficient of Determination

$$r^2 = \frac{\left(\sum xy\right)^2}{\sum x^2 \sum y^2}$$

note: similar formula as for r^2 in regression, but different interpretation
= variability in one variable (either X or Y) accounted for by correlating w/ 2nd variable

1.4 Standard Error of Correlation Coefficient

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

2 Testing significance of Correlation Coefficient

2.1 H_0: $\rho = 0$ H_A: $\rho \ne 0$

- to test H_0: $\rho = \rho_0$ H_A: $\rho \ne \rho_0$ (where $\rho_0 \ne 0$)
  must use different method (Fisher's z): see Zar § 19.2, pp.381-383

2.2 t-test: $t = \frac{r}{s_r}$

reject H_0 if $|t| \ge t_{\alpha(2),\nu}$ where $\nu = n - 2$

2.3 F-test: $F = \frac{1+|r|}{1-|r|}$

reject H_0 if $F \ge F_{\alpha(2),\nu,\nu}$

2.4 Assumptions for testing significance (no assumptions required to calculate r)

2.4.1 X & Y values sampled at random from normally distributed populations
2.4.2 Y, X bivariate normal distribution