FORMULA SHEET FOR
MID-TERM

BASIC STATISTICS:

1. Var(X) = $E(X-\mu_x)^2$

2. Cov(X,Y) = $E(X-\mu_x)(Y-\mu_y)$; $Corr(X,Y) = Cov(X,Y)/(Var(X)\cdot Var(Y))^{1/2}$

3. E(aX + bY) = aE(X) + bE(Y)

4. Var(aX + bY) = $a^2Var(X) + b^2Var(Y) + 2Cov(X,Y)$

5. Sample Mean: $\bar{Y} = \sum_1^N Y_i$

6. Sample Variance: $s^2 = \sum_1^N (Y_i - \bar{Y})^2 /(N-1)$

7. t-statistic for testing population mean:

$$t_{N-1} = \frac{\bar{Y} - \mu_{Y,0}}{se(\bar{Y})}; \text{ where } se(\bar{Y}) = s/\sqrt{N}$$

8. $(1-\alpha)\%$ confidence interval for μ

$$\Pr(\bar{Y} - t_{N-1,\alpha/2}\cdot se(\bar{Y}) < \mu < \bar{Y} + t_{N-1,\alpha/2}\cdot se(\bar{Y})) = 1-\alpha$$

9. Approximate t-statistic for testing difference in means (variances assumed unequal):

$$t = \frac{\bar{Y}_1 - \bar{Y}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \to Z = N(0,1)$$

10. Exact t-statistic for testing difference in means (variances assumed equal):

$$t = \frac{\bar{Y}_1 - \bar{Y}_2}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \to t_{n_1+n_2-2}$$

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

11. F-Test for equal variances across two populations

$$F_{v_1, v_2} = \frac{s_1^2}{s_2^2}$$

where $v_1 = n_1 - 1$ and $v_2 = n_2 - 1$. Also, $F_{1-\alpha/2}(v_1, v_2) = \dfrac{1}{F_{\alpha/2}(v_1, v_2)}$

SOME OLS REGRESSION FORMULAS:

12. $\hat{\beta}_0 = \overline{Y} - \hat{\beta}_1 \overline{X}$; $\qquad Var(\hat{\beta}_0) = \dfrac{\sigma^2 \sum_1^N X_i}{N \sum_1^N (X_i - \overline{X})^2}$

13. $\hat{\beta}_1 = \dfrac{\sum_1^N (X_i - \overline{X}) Y_i}{\sum_1^2 (X_i - \overline{X})^2} = \sum_1^N w_i Y_i$; $\qquad Var(\hat{\beta}_1) = \dfrac{\sigma^2}{\sum_1^N (X_i - \overline{X})^2}$

14. $\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X$

15. TSS = ESS + SSR ; $\sum_1^N (Y_i - \overline{Y})^2 = \sum_1^N (\hat{Y}_i - \overline{Y})^2 + \sum_1^N (Y_i - \hat{Y}_i)^2$; $R^2 = \dfrac{ESS}{TSS}$

16. $t = \dfrac{\hat{\beta}_i - \beta_i}{se(\hat{\beta}_i)}$

17. one-tailed p-value: $\Pr(t_0 < t)$ or $\Pr(t < t_0)$

18. two-tailed p-value: $\Pr(|t_0| < t)$

19. $\Pr(\hat{\beta}_i - t_{N-K, \alpha/2} \cdot se(\hat{\beta}_i) < \beta_i < \hat{\beta} + t_{N-K, \alpha/2} \cdot se(\hat{\beta}_i)) = 1 - \alpha$

20. $F_{overall} = \dfrac{R^2 / (K-1)}{(1 - R^2)/(N-K)}$

21. $F = \dfrac{(RSS_R - RSS_U)/J}{RSS_U / (N-K)} = \dfrac{(R_U^2 - R_R^2)/J}{(1 - R_U^2)/(N-K)}$