Formula Sheet (Page 1)

1. Addition Rule (events not mutually exclusive): $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

2. Addition Rule (events not mutually exclusive):
$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$

3. Addition Rule (mutually exclusive events): $P(A \cup B) = P(A) + P(B)$

4. Multiplication Rule (dependent events): $P(A \cap B) = P(A)P(B|A)$

5. Multiplication Rule (independent events): $P(A \cap B) = P(A)P(B)$

6. Conditional Probability: $P(B|A) = \frac{P(A \cap B)}{P(A)}$

7. Binomial: $f(x) = \binom{n}{x} p^x (1-p)^{n-x},\ x = 0, 1, \ldots, n;\quad \binom{n}{x} = \frac{n!}{x!(n-x)!};\quad \mu = np,\ \sigma^2 = np(1-p)$

8. Negative Binomial: $f(x) = \binom{x-1}{r-1}(1-p)^{x-r}p^r,\ x = r, r+1, r+2, \ldots;\quad \mu = \frac{r}{p}, \sigma^2 = \frac{r(1-p)}{p^2}$

9. Hypergeometric Distribution: $f(x) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}},\quad \mu = np,\ \sigma^2 = np(1-p)\left(\frac{N-n}{N-1}\right)$

10. Poisson Distribution: $f(x) = \frac{e^{-\lambda}\lambda^x}{x!}, x = 0, 1, 2, \ldots;\qquad \mu = \sigma^2 = \lambda$

11. Normal Distribution: $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2},\ -\infty < x < \infty$

12. Exponential Distribution: $f(x) = \lambda e^{-\lambda x},\ 0 < x < \infty;\ \mu = \frac{1}{\lambda},\ \sigma^2 = \frac{1}{\lambda^2}$

13. Marginal Distributions: $f_X(x) = \int f_{XY}(x,y)\,dy,\ f_Y(y) = \int f_{XY}(x,y)\,dx$

14. Correlation: $\rho_{XY} = \frac{\mathrm{cov}(X,Y)}{\sqrt{V(X)V(Y)}} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$

15. Transformation to standard normal: $z = \frac{X-\mu}{\sigma}$

16. Sample variance: $s^2 = \frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n-1} = \frac{\sum_{i=1}^{n} x_i^2 - n\overline{x}^2}{n-1}$

17. Outliers: $Q_1 - 1.5\,\mathrm{IQR}, Q_3 + 1.5\,\mathrm{IQR}$

18. Normal probability plot: $\Phi(z_j) = \frac{j-0.5}{n},\ \ j = 1, 2, \ldots, n$

19. Central Limit Theorem formula: $z = \frac{\overline{x}-\mu}{\sigma/\sqrt{n}}$

20. z confidence interval for the mean: $\overline{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$

Formula Sheet (Page 2)

21. t confidence interval for the mean: $\overline{x} \pm t_{\alpha/2,n-1}\frac{s}{\sqrt{n}}$

22. Confidence interval for a proportion: $\widehat{p} \pm z_{\alpha/2}\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}}$

23. z test for a mean: $Z_0 = \frac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}$; $\quad \beta = \Phi\left(z_{\alpha/2} - \frac{\delta\sqrt{n}}{\sigma}\right) - \Phi\left(-z_{\alpha/2} - \frac{\delta\sqrt{n}}{\sigma}\right)$

24. t test for a mean: $T_0 = \frac{\overline{x}-\mu_0}{s/\sqrt{n}}$ **25.** z test for proportions: $Z_0 = \frac{X-np_0}{\sqrt{np_0(1-p_0)}} = \frac{\widehat{p}-p_0}{\sqrt{p_0(1-p_0)/n}}$;

$$\beta = \Phi\left(\frac{p_0-p+z_{\alpha/2}\sqrt{p_0(1-p_0)/n}}{\sqrt{p(1-p)/n}}\right) - \Phi\left(\frac{p_0-p-z_{\alpha/2}\sqrt{p_0(1-p_0)/n}}{\sqrt{p(1-p)/n}}\right)$$

Confidence interval for a difference in means:

26. Variances equal: $\overline{x}_1 - \overline{x}_2 \pm t_{\alpha/2,n_1+n_2-2}s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$, $\quad s_p^2 = \frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}$

27. Variances unequal: $\overline{x}_1 - \overline{x}_2 \pm t_{\alpha/2,\nu}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$, $\quad \nu = \left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)^2 \Big/ \left(\frac{(s_1^2/n_1)^2}{n_1-1}+\frac{(s_2^2/n_2)^2}{n_2-1}\right)$

28. t test for comparing two means (variances equal): $t_{n_1+n_2-2} = \frac{\overline{x}_1-\overline{x}_2}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$

29. t test for comparing two means (variances unequal): $t_\nu = \frac{\overline{x}_1-\overline{x}_2}{\sqrt{(s_1^2/n_1)+(s_2^2/n_2)}}$,

30. Single variable Least Squares Regression line: $\widehat{\beta}_0 = \overline{y} - \widehat{\beta}_1\overline{x}$, $\widehat{\beta}_1 = \frac{s_{xy}}{s_{xx}}$, where

$$s_{xx} = \sum_{i=1}^{n} x_i^2 - n\overline{x}^2 \text{ and } s_{xy} = \sum_{i=1}^{n} x_i y_i - n\overline{x}\,\overline{y}$$

31. t-test for single variable regression: $t_{n-2} = \frac{\widehat{\beta}_1}{\sqrt{\widehat{\sigma}^2/s_{xx}}}$, where $\widehat{\sigma}^2 = \frac{SS_E}{n-2}$

32. Residual sum of squares: $SS_E = \sum_{i=1}^{n}(y_i - \widehat{y}_i)^2$ **33.** Regression sum of squares: $SS_R = \widehat{\beta}_1 s_{xy}$

34. Total sum of squares: $SS_T = \sum_{i=1}^{n} y_i^2 - n\overline{y}^2$

35. Prediction Interval: $\widehat{y}_0 \pm t_{\alpha/2,n-2}\sqrt{\widehat{\sigma}^2\left(1+\frac{1}{n}+\frac{(x_0-\overline{x})^2}{s_{xx}}\right)}$

36. Sample correlation coefficient: $r = \frac{s_{xy}}{\sqrt{s_{xx}s_{yy}}}$ **37.** Coefficient of determination: $R^2 = \frac{SS_R}{SS_T}$

38. Total sum of squares: $SS_T = \sum_{i=1}^{a}\sum_{j=1}^{n_i}(y_{ij} - \overline{y}_{\bullet\bullet})^2 = \sum_{i=1}^{a}\sum_{j=1}^{n_i} y_{ij}^2 - \frac{y_{\bullet\bullet}^2}{N}$ (d.f. $= N-1$)

39. Error sum of squares: $SS_E = \sum_{i=1}^{a}\sum_{j=1}^{n_i}(y_{ij} - \overline{y}_{i\bullet})^2 = \sum_{i=1}^{a}(n_i - 1)s_i^2$ (d.f. $= N-a$)

40. Treatment sum of squares: $SS_{\text{Treatments}} = \sum_{i=1}^{a} n_i(\overline{y}_{i\bullet} - \overline{y}_{\bullet\bullet})^2 = \sum_{i=1}^{a}\frac{y_{i\bullet}^2}{n_i} - \frac{y_{\bullet\bullet}^2}{N}$ (d.f. $= a-1$)

41. Fisher's LSD Test: $\text{LSD} = t_{\alpha/2,N-a}\sqrt{MSE\left(\frac{1}{n_i}+\frac{1}{n_j}\right)}$, **42.** Fisher's CI: $\overline{y}_{i\bullet} - \overline{y}_{j\bullet} \pm \text{LSD}$