MAT 2379, Introduction to Biostatistics

Midterm Formula Sheet

- Addition Rule: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- Conditional probability of A given B:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

- Total probability rule: $P(A) = P(A \cap B) + P(A \cap B') = P(A|B)P(B) + P(A|B')P(B')$
- Bayes' rule

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

- Diagnostic tests:

$$\begin{aligned}
\text{false-positive-rate} &= P(\text{Test}+|\text{True}-) \\
\text{false-negative-rate} &= P(\text{Test}-|\text{True}+) \\
\text{specificity} &= P(\text{Test}-|\text{True}-) \\
\text{sensitivity} &= P(\text{Test}+|\text{True}+) \\
\text{positive predictive value} &= P(\text{True}+|\text{Test}+) \\
\text{negative predictive value} &= P(\text{True}-|\text{Test}-)
\end{aligned}$$

- Events A and B are independent if $P(A \cap B) = P(A)P(B)$
- Expected value of a discrete random variable X:

$$\mu = E(X) = \sum_x xf(x), \quad \text{where} \quad f(x) = P(X = x)$$

- Variance of a discrete random variable X:

$$\sigma^2 = \text{Var}(X) = \sum_x (x-\mu)^2 f(x) = \sum_x x^2 f(x) - \mu^2, \quad \text{where} \quad f(x) = P(X = x)$$

- Cumulative distribution function of a random variable X: $F(x) = P(X \leq x)$
- Standardization: If X is a normal random variable with mean μ and variance σ^2, then

$$Z = \frac{X - \mu}{\sigma} \quad \text{has a standard normal distribution}$$

- Sample mean of the observations $x_1, \ldots, x_n$:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

- Sample variance of the observations $x_1, \ldots, x_n$:

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2 = \frac{1}{n-1} \left(\sum_{i=1}^{n} x_i^2 - n\bar{x}^2 \right) = \frac{1}{n-1} \left[\sum_{i=1}^{n} x_i^2 - \frac{1}{n} \left(\sum_{i=1}^{n} x_i \right)^2 \right]$$

- To calculate the quartiles, we arrange the data $x_1, x_2 \ldots, x_n$ in ascending order:

$$y_1 \leq y_2 \leq \ldots \leq y_n$$

a) To compute the first quartile, we write

$$\frac{n+1}{4} = r + \frac{a}{4}, \quad r = \text{integer}, a = 0, 1, 2 \text{ or } 3.$$

The first quartile is:

$$q_1 = \begin{cases} y_r, & \text{if } (n+1)/4 = r \\ (0.75)\, y_r + (0.25)\, y_{r+1}, & \text{if } (n+1)/4 = r + 1/4 \\ (0.5)\, y_r + (0.5)\, y_{r+1}, & \text{if } (n+1)/4 = r + 2/4 \\ (0.25)\, y_r + (0.75)\, y_{r+1}, & \text{if } (n+1)/4 = r + 3/4 \end{cases}$$

b) To compute the third quartile, we write

$$\frac{3(n+1)}{4} = r + \frac{a}{4}, \quad r = \text{integer}, a = 0, 1, 2 \text{ or } 3.$$

The third quartile is:

$$q_3 = \begin{cases} y_r, & \text{if } (3(n+1)/4 = r \\ (0.75)\, y_r + (0.25)\, y_{r+1}, & \text{if } 3(n+1)/4 = r + 1/4 \\ (0.5)\, y_r + (0.5)\, y_{r+1}, & \text{if } 3(n+1)/4 = r + 2/4 \\ (0.25)\, y_r + (0.75)\, y_{r+1}, & \text{if } 3(n+1)/4 = r + 3/4 \end{cases}$$