Review of Introduction to mathematical statistics

Contents you should know before you take this class

- Mathematical probability and random variables
- Discrete and continuous random variables
- Probability mass (density) function and cumulative distribution function
- Mathematical expecation, variance, and moment generating functions
- Variable transformation technique
- Special discrete random variables: Binomial, Poisson, Geometric, Negative binomial
- Special continuous random variables: Exponential, Gamma, Normal, χ^2, Beta, F, Uniform, t-distributions
- For above special distritbuions, I will assume that you are able to compute and derive expecation, variance, moment generating functions. It is also important to know the relationship of distributions.

Review of order statistics (Ch 4.4)

- Let X_1 and X_2 be independent random variables with common continuous pdf $f_X(\cdot)$. Suppose that we want to find the pdf of $Y_1 = \min(X_1, X_2)$.

  $F_{Y_1}(y) =$

- Suppose that we want to find the pdf of $Y_2 = \max(X_1, X_2)$.

$F_{Y_2}(y) =$

- Now, suppose that we have three independent random variables X_1, X_2, X_3 having common pdf $f_X(\cdot)$. Let $Y_1 < Y_2 < Y_3$ be the order statistics from X_1, X_2, X_3. How to find the pdf of each order statistics?

- In general, the pdf of jth order statistic Y_j from $X_1, \ldots, X_n$ is

$$f_{Y_j}(y) = \frac{n!}{(j-1)!(n-j)!} \left(F_X(y)\right)^{j-1} f_X(y) \left(1 - F_X(y)\right)^{n-j}$$

- More generally the joint pdf of Y_i and Y_j is

$$\begin{aligned} f_{Y_i, Y_j}(y_i, y_j) =& \frac{n!}{(i-1)!(j-i-1)!(n-j)!} f_X(y_i) f_X(y_j) \\ & \times \left(F_X(y_i)\right)^{i-1} \left[F_X(y_j) - F_X(y_i)\right]^{j-i-1} \left[1 - F_X(y_j)\right]^{n-j} \end{aligned}$$

- Example: $X_1, X_2, X_3, X_4 \overset{iid}{\sim} f(x) = 2x, 0 < x < 1.$

- For more exercises, see exercises 4.4.5, 4.4.6, 4.4.8, 4.4.9 in the text.

- Heuristic derivation

Suppose that $X_1, \ldots, X_n$ is a random sample from $f_X(x)$. Let us consider $P(y < Y_j \leq y + \epsilon)$:

$$\begin{aligned} P(y < Y_j \leq y + \epsilon) = P[&(j-1)X_i'\text{s are less than } y\text{, one } X_i \text{ is} \\ &\text{between } y \text{ and } y+\epsilon\text{, and } (n-j-1)X_i's \\ &\text{are greater than } y] \end{aligned}$$

For each X_i, there are three possible cases:

1. $X_i \leq y$ with probability $F_X(y)$
2. $y < X_i \leq y + \epsilon$ with probability $F_X(y + \epsilon) - F_X(y)$
3. $X_i > y$ with probability $1 - F_X(y + \epsilon)$

$$P(y < Y_j \leq y + \epsilon) = \frac{n!}{(j-1)!1!(n-j)!} \left(F_X(y)\right)^{j-1} f_X(y) \left(1 - F_X(y)\right)^{n-j}$$

$$
\begin{aligned}
f_{Y_i,Y_j}(y_i,y_j) =& \frac{n!}{(i-1)!1!(j-i-1)!1!(n-j)!} \left(F_X(y_i)\right)^{i-1} f_X(y_i) \\
& \times \left[F_X(y_j) - F_X(y_i)\right]^{j-i-1} f_X(y_j) \left[1 - F_X(y_j)\right]^{n-j}
\end{aligned}
$$