Intro Prob Lecture Notes

William Sun (Partially based on Joon Hyuck Choi's Notes)

April 7, 2017

Order Statistics

- Suppose $X_1, X_2, \ldots, X_n$ are independent continuous random variables all having the same pdf $f(x)$. That is, $X_1, X_2, \ldots, X_n \sim$ i.i.d. F
- We define for each $j = 1, 2, \ldots n$:
  - $Y_j = j$th smallest value among $X_1, X_2, \ldots, X_n$
    - (where $Y_1 \leq Y_2 \leq \cdots \leq Y_n$)
- First, we look at $Y_n = max\{X_1, X_2, \ldots, X_n\}$. What's its pdf?
  - Let's use the cdf method:
  -

$$\begin{aligned}
F_{Y_n}(y) &= P(max\{X_1, X_2, \ldots, X_n\} \leq y) \\
&= P(X_1 \leq y, X_2 \leq y, \ldots, X_n \leq y) \\
&= P(X_1 \leq y)P(X_2 \leq y) \ldots P(X_n \leq y) \\
&= (F(y))^n \text{ where } F(y) = P(X_1 \leq y)
\end{aligned}$$

  - So,

$$f_{Y_n}(y) = n(F(y))^{n-1} \cdot f(y)$$

-

How about Y_1?

$$F_{Y_1}(y) = p(min\{X_1, X_2, \dots, X_n\} \leq y)$$
$$= 1 - p(min\{X_1, X_2, \dots, X_n\} > y)$$
$$= 1 - p(X_1 > y, X_2 > y, \dots, X_n > y)$$
$$= 1 - (p(X_1 > y))^n$$
$$= 1 - (1 - F(y))^n$$

- So,

$$f_{Y_1}(y) = n(1 - F(y))^{n-1} f(y)$$

- Ex: (Grocery Clerks) $X_1, X_2, X_3 \sim$ independent exp(1)
  - pdf $f(x) = e^{-x}$ for $x > 0$
  - cdf $F(x) = 1 - e^{-x}$ for $x > 0$
  - $p(min\{X_2, X_2, X_3\} > t) = \int\limits_t^\infty 3e^{-3x} dx = e^{-3t}$
- Now let's look at Y_j for $X_1, X_2, X_3, \dots X_n$ *sim* i.i.d. F
  - Distribution of Y_j?
    - *

$$f_{Y_j} = \lim_{h \to 0} \frac{P(y - h < Y_J \leq y)}{h}$$

      · $0 < h << 1$ (h is much much less than 1, very close to 0)

      ·

$$P(y - h < Y_j \leq y) = \binom{n}{j-1} P(X_1 \leq y - h)^{j-1} \cdot \binom{n-j+1}{1} P(y - h < X_1 \leq y) \cdot P(X_1 > y)^{n-j}$$
$$= \binom{n}{j-1} F(y-h)^{j-1} \frac{F(y) - F(y-h)}{h} (n - j + 1) \cdot (1 - F(y))^{n-j}$$
$$f_{Y_j} = \frac{n!}{(j-1)!(n-j)!} F(y)^{j-1} f(y) (1 - F(y))^{n-j}$$

- Ex: $U_1, U_2, \dots, U_n \sim$ i.i.d. uniform(0, 1)
  - What is the pdf of Y_j?

–

$$
\begin{aligned}
f_{Y_j}(y) &= \frac{n!}{(j-1)!(n-j)!} y^{j-1} \cdot 1 \cdot (1-y)^{n-j} \sim \text{Beta}(j, n-j+1) \\
&= \frac{\Gamma(n+1)}{\Gamma(j)\Gamma(n-j+1)} y^{j-1} (1-y)^{(n-j+1)-1}
\end{aligned}
$$

- Exercise $X_1, X_2, \ldots X_{2n-1} \sim$ iid Uniform(0, 1)
  - Find the pdf of Y_n (this is the median)
    - Has maximum density at $y = \frac{1}{2}(\frac{1}{2} = mode)$