Stat 562
1-29-19

One last relationship between distributions:

①

Recall the Beta distribution (11/13/18 notes).

If $F \sim F_{r_1, r_2}$, then

$$\frac{\frac{r_1}{r_2}F}{1+\frac{r_1}{r_2}F} \sim \text{Beta}\left(\frac{r_1}{2}, \frac{r_2}{2}\right)$$

②

Order statistics

Let $X_1, X_2, \ldots, X_n \sim$ iid with cdf $F(x)$

Let $X_{(1)} = \min\{X_1, \ldots, X_n\}$

$X_{(2)}$ = second smallest of $\{X_1, \ldots, X_n\}$

$\vdots$ etc.

$X_{(n)} = \max\{X_1, \ldots, X_n\}$

These are called the order statistics

Note $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$

③

Find the distribution of $X_{(1)}$.

$$F_1(x) = P[X_{(1)} \le x) = 1 - P(X_{(1)} > x)$$

$$= 1 - P(X_i > x \ \forall i)$$

$$= 1 - \prod_{i=1}^{n} P(X_i > x) \quad \text{by indep}$$

$$= 1 - (1 - F(x))^n$$

In the continuous case,

$$f_1(x) = -n(1 - F(x))^{n-1}(-f(x))$$

④

$$f_1(x) = n(1 - F(x))^{n-1} f(x)$$

Find the distribution of $X_{(n)}$.

$$F_n(x) = P(X_{(n)} \le x)$$

$$= P(X_i \le x \ \forall i)$$

$$= \prod_{i=1}^{n} P(X_i \le x) = (F(x))^n$$

In the continuous case,

$$f_n(x) = n(F(x))^{n-1} f(x)$$

Find the distribution of $X_{(k)}$ ⑤

$$F_k(x) = P\left(X_{(k)} \le x\right)$$
$$= P\left(X_{(k)} \le x \cap X_{(k+1)} \le x\right)$$
$$+ \; P\left(X_{(k)} \le x \cap X_{(k+1)} > x\right)$$
$$= P\left(X_{(k+1)} \le x\right) + P\left(X_{(k)} \le x \cap X_{(k+1)} > x\right)$$
$$= F_{k+1}(x) + P\left(\text{exactly } k \text{ of } \{X_1, \dots, X_n\} \text{ are } \le x\right)$$

⑥

$$F_k(x) = F_{k+1}(x) + \binom{n}{k}\left(F(x)\right)^k \left(1 - F(x)\right)^{n-k}$$

$$F_k(x) - F_{k+1}(x) = \binom{n}{k}\left(F(x)\right)^k \left(1 - F(x)\right)^{n-k}$$

In the continuous case,

$$f_k(x) - f_{k+1}(x) = \binom{n}{k}\Big[\left(F(x)\right)^k (n-k)\left(1 - F(x)\right)^{n-k-1}\left(-f(x)\right) + \left(1 - F(x)\right)^{n-k} k \left(F(x)\right)^{k-1} f(x)\Big]$$

$$f_k(x) - f_{k+1}(x)$$
$$= k\binom{n}{k}(F(x))^{k-1}(1-F(x))^{n-k} f(x)$$
$$- (n-k)\binom{n}{k}(F(x))^{k}(1-F(x))^{n-k-1} f(x)$$

$$= k\binom{n}{k}(F(x))^{k-1}(1-F(x))^{n-k} f(x)$$
$$-(k+1)\binom{n}{k+1}(F(x))^{k}(1-F(x))^{n-(k+1)} f(x)$$
$$= g_k(x) - g_{k+1}(x)$$

$$(n-k)\binom{n}{k} = (n-k)\frac{n!}{k!(n-k)!}$$
$$= \frac{n!}{k!(n-k-1)!}$$
$$= \frac{(k+1)\, n!}{(k+1)!(n-k-1)!} = (k+1)\binom{n}{k+1}$$

We would like to show that $f_k(x) = g_k(x)\ \forall k$ 8

$k=1$: $f_1(x) = n(1-F(x))^{n-1} f(x)$

$$g_1(x) = 1\binom{n}{1}(F(x))^0(1-F(x))^{n-1} f(x)$$

So $f_1(x) = g_1(x)$

Now $f_1(x) - f_2(x) = g_1(x) - g_2(x)$

$$\therefore f_2(x) = g_2(x)$$

Repeating this, we get our result.

(9)

$$\therefore\ f_k(x) = k\binom{n}{k}[F(x)]^{k-1}[1-F(x)]^{n-k} f(x)$$

Example: Suppose $X_1, \dots, X_n \sim$ iid $Exp(\lambda)$

$$\left(\begin{array}{l} f(x) = \lambda e^{-\lambda x} \quad x>0 \\ F(x) = 1-e^{-\lambda x} \end{array}\right)$$

Find the density of the k^{th} order statistic

$$f_k(x) = k\binom{n}{k}\left(1-e^{-\lambda x}\right)^{k-1}\left(e^{-\lambda x}\right)^{n-k}\lambda e^{-\lambda x}$$

(10)

Special case: $k=1$

$$f_1(x) = \lambda \cdot 1\binom{n}{1}\left(e^{-\lambda x}\right)^{n-1} e^{-\lambda x}$$

$$= n\lambda e^{-n\lambda x} \quad \sim Exp(n\lambda)$$

Special case: $k=n$

$$f_n(x) = n\binom{n}{n}\left(1-e^{-\lambda x}\right)^{n-1} f(x)$$

$$= n\lambda e^{-\lambda x}\left(1-e^{-\lambda x}\right)^{n-1}$$

Next big task:

Find the joint pdf of any 2 order statistics

$X_{(i)}$ & $X_{(j)}$

$$f_{ij}(x_i, x_j) = \frac{\partial^2}{\partial x_i \partial x_j} F_{ij}(x_i, x_j)$$

$$= \frac{\partial}{\partial x_j}\left[\frac{\partial}{\partial x_i} F_{ij}(x_i, x_j)\right]$$

$$= \lim_{\Delta_j \to 0} \frac{\frac{\partial}{\partial x_i} F_{ij}(x_i, x_j+\Delta_j) - \frac{\partial}{\partial x_i} F_{ij}(x_i, x_j)}{\Delta_j}$$

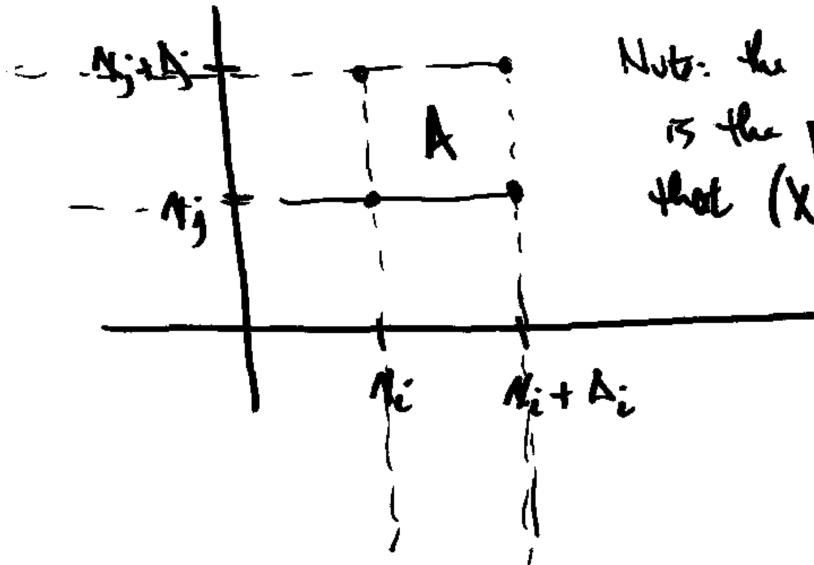
(12)

$$= \lim_{\Delta_j \to 0} \frac{\lim_{\Delta_i \to 0} \frac{F_{ij}(x_i+\Delta_i, x_j+\Delta_j) - F_{ij}(x_i, x_j+\Delta_j)}{\Delta_i} - *}{\Delta_j}$$

$$\text{where } * = \lim_{\Delta_i \to 0} \frac{F_{ij}(x_i+\Delta_i, x_j) - F_{ij}(x_i, x_j)}{\Delta_i}$$

$$= \lim_{\substack{\Delta_i \to 0 \\ \Delta_j \to 0}} \frac{F_{ij}(x_i+\Delta_i, x_j+\Delta_j) - F_{ij}(x_i, x_j+\Delta_j) - F_{ij}(x_i+\Delta_i, x_j) + F_{ij}(x_i, x_j)}{\Delta_i \Delta_j}$$

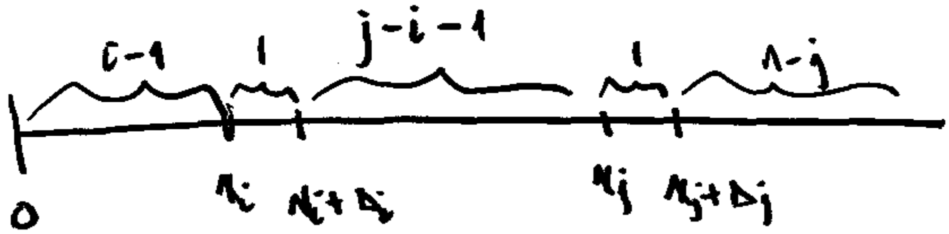
(13)



Note: the numerator is the probability that $(X_{(i)}, X_{(j)})$ lies in box A

$$\lim_{\substack{\Delta_i \to 0 \\ \Delta_j \to 0}} \frac{P\left(x_i < X_{(i)} < x_i+\Delta_i \;\cap\; x_j < X_{(j)} < x_j+\Delta_j\right)}{\Delta_i \Delta_j}$$

(14)



$$n-(i-1+1+1+n-j)=j-i-1$$

To be continued