Lecture 12

Order statistics

Maximum and minimum

Several random variables

Random permutations

Bivariate distributions

Objects of interest

Joint probabilities	$P\{X = x, Y = y\}$	$p_{XY}(x, y)$
Marginal probs	$P\{X = x\}$	$p_X(x) = \sum_y p_{XY}(x, y)$
	$P\{Y = y\}$	$p_Y(y) = \sum_x p_{XY}(x, y)$
Conditional probs	$P\{X = x \mid Y = y\}$	$p_{X\mid Y}(x\mid y) = \frac{p_{XY}(x,y)}{p_Y(y)}$
	$P\{Y = y \mid X = x\}$	$p_{Y\mid X}(y\mid x) = \frac{p_{XY}(x,y)}{p_X(x)}$

Order statistics

- Data: $\{X_1, X_2, \ldots, X_n\}$
- Order statistics: $\{X_{(1)} \leq X_{(2)} \leq \ldots \leq X_{(n)}\}$

Example

$$\begin{aligned} \{X_1, X_2, X_3, X_4, X_5\} &= \{7, 5, 3, 8, 5\} \\ \{X_{(1)}, X_{(2)}, X_{(3)}, X_{(4)}, X_{(5)}\} &= \{3, 5, 5, 7, 8\} \end{aligned}$$

- If the data is random, then the order statistics $\{X_{(1)}, X_{(2)}, \ldots, X_{(n)}\}$ are a set of n random variables.
- Even if the data is made of independent observations, then the order statistics are dependent.

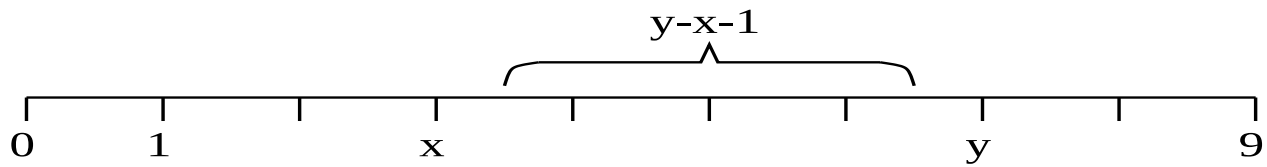
Maximum $= X_{(n)}$, Minimum $= X_{(1)}$

Example: Choose 3 numbers at random, without replacement, from $\{0, 1, \ldots, 9\}$. Set

$$X \equiv \text{minimum} = X_{(1)}, \qquad Y \equiv \text{maximum} = X_{(3)}.$$

- Range of X : $\{0, 1, \ldots, 7\}$.
- Range of Y : $\{2, 3, \ldots, 9\}$.
- Range of (X, Y) : $0 \leq x \leq y - 2 \leq 7$.
- Number of possible triples: $\binom{10}{3} = \frac{10!}{3!7!} = \frac{10 \cdot 9 \cdot 8}{3 \cdot 2 \cdot 1} = 120$.
- All ordered triples are equally likely.

$$\begin{aligned} p_{XY}(x, y) &= P\{X = x, Y = y\} \\ &= \frac{\text{Number of samples with Max}{=}y \text{ and Min}{=}x}{120} \end{aligned}$$



$$p_{XY}(x,y) = \begin{cases} \frac{(y-x-1)}{120} & \text{if } 0 \le x \le y-2 \le 7\ , \\ 0 & \text{otherwise.} \end{cases}$$

$$\begin{aligned} p_X(x) &= \sum_y p_{XY}(x,y) \\ &= \sum_{y=x+2}^{9} \frac{(y-x-1)}{120} \\ &= (1+2+\ldots+(8-x))\,/120 \\ &= (8-x)(9-x)/240, \qquad 0 \le x \le 7 \end{aligned}$$

x	0	1	2	3	4	5	6	7
$p_X(x)$	$\frac{36}{120}$	$\frac{28}{120}$	$\frac{21}{120}$	$\frac{15}{120}$	$\frac{10}{120}$	$\frac{6}{120}$	$\frac{3}{120}$	$\frac{1}{120}$

$$
\begin{aligned}
p_Y(y) &= \sum_x p_{XY}(x,y) \\
&= \sum_{x=0}^{y-2} \frac{(y-x-1)}{120} \\
&= ((y-1)+(y-2)+\ldots+1)/120 \\
&= y(y-1)/240, \qquad\qquad 2 \le y \le 9
\end{aligned}
$$

y	2	3	4	5	6	7	8	9
$p_Y(y)$	$\frac{1}{120}$	$\frac{3}{120}$	$\frac{6}{120}$	$\frac{10}{120}$	$\frac{15}{120}$	$\frac{21}{120}$	$\frac{28}{120}$	$\frac{36}{120}$

$$
\begin{aligned}
p_{X|Y}(x|y) = \frac{p_{XY}(x,y)}{p_Y(y)} &= \frac{(y-x-1)/120}{y(y-1)/240} \\
&= \frac{2(y-x-1)}{y(y-1)} \qquad\qquad 0 \le x \le y-2
\end{aligned}
$$

$$p_{X|Y}(x|y) = \frac{2(y-x-1)}{y(y-1)} \qquad 0 \le x \le y-2$$

x	0	1	2	3	4	5	6	7
$p_{X\|Y}(x\|3)$	$\frac{2}{3}$	$\frac{1}{3}$						
$p_{X\|Y}(x\|5)$	$\frac{4}{10}$	$\frac{3}{10}$	$\frac{2}{10}$	$\frac{1}{10}$				
$p_{X\|Y}(x\|9)$	$\frac{8}{36}$	$\frac{7}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

- Note:

$$\frac{2(y-x-1)}{y(y-1)} = \frac{y-x-1}{\binom{y}{2}}$$

Several random variables
(X, Y, Z), $(X_1, X_2, \ldots, X_n)$

1: Order statistics, $(X, Y, Z) = (X_{(1)}, X_{(2)}, X_{(3)})$.

- Same example as before:

$$P\{X = x, Y = y, Z = z\}$$
$$= p_{XYZ}(x, y, z) = \begin{cases} \frac{1}{120} & \text{if } 0 \le x < y < z \le 9 \ , \\ 0 & \text{otherwise.} \end{cases}$$

- Same (X, Y, Z), but out of a sample of 5:

$$p_{XYZ}(x, y, z) = \begin{cases} \frac{\binom{9-z}{2}}{\binom{10}{5}} & \text{if } 0 \le x < y < z \le 7 \ , \\ 0 & \text{otherwise.} \end{cases}$$

2: Random permutations

- $\hat{X} \equiv \{X_1, X_2, \ldots, X_n\}$, a random permutation of $\{1, 2, \ldots, n\}$.

$$n = 5: \qquad \hat{X} = \{3, 5, 4, 1, 2\}, \quad \hat{X} = \{4, 1, 2, 3, 5\}, \ \ldots$$

$$p_{\hat{X}}(\hat{x}) = p_{(X_1, X_2, \ldots, X_n)}(x_1, x_2, \ldots, x_n) = \begin{cases} \frac{1}{n!} \\ 0 \end{cases}$$

- Examples: Card deals, Drawing (all) Lotto numbers, ...

Properties of several RV's

1: Independence $X_1, X_2, \ldots, X_n$ are independent if, and only if,

$$p_{(X_1,X_2,\ldots,X_n)}(x_1, x_2, \ldots, x_n) \;= p_{X_1}(x_1)p_{X_2}(x_2)\cdots p_{X_n}(x_n)$$

for every $(x_1, x_2, \ldots, x_n)$.

- Note: Independence in pairs is not enough!

  - $H_1 = \{\text{Head on first toss}\} = \{HH, HT\} \Rightarrow \; P\{H_1\} = 1/2$
  - $H_2 = \{\text{Head on 2nd toss}\} = \{TH, HH\} \Rightarrow P\{H_2\} = 1/2$
  - $S = \{\text{Both tosses the same}\} = \{HH, TT\} \Rightarrow \; P\{S\} = 1/2$

$$X_1 = 1_{H_1}, \quad X_2 = 1_{H_2}, \quad X_3 = 1_S$$

where

$$1_A \;=\; 1, \text{ if } A \text{ occurs}, \;=\; 0, \text{ otherwise.}$$

2: Equivalent formulation of independence

$$P\{X_1 \in A_1, X_2 \in A_2, \ldots, X_n \in A_n\}$$
$$= \; P\{X_1 \in A_1\}P\{X_2 \in A_2\} \cdots P\{X_n \in A_n\},$$

for all sets $A_1, A_2, \ldots, A_n$.

3: Functions of independent random variables are independent.

$$X_1, X_2, \ldots, X_n \text{ independent}$$
$$\Rightarrow \; f_1(X_1), f_2(X_2), \ldots, f_n(X_n) \text{ independent.}$$

4: Disjoint blocks of IRV's are independent.

$$X_1, X_2, \ldots, X_n \text{ independent}$$
$$\Rightarrow (X_1, \ldots, X_{k_1}), \ldots \left(X_{k_{m-1}}, \ldots, X_{k_m}\right) \text{ independent.}$$

5: Functions of disjoint blocks of IRV's are independent.

6: One-dimensional marginals

$$p_{X_1}(x) = \sum_{x_2} \cdots \sum_{x_n} p_{X_1,X_2,\ldots,X_n}(x, x_2, \ldots, x_n).$$

7: k-dimensional marginals

$$p_{X_1,\ldots,X_k}(x_1, \ldots, x_k) = \sum_{x_{k+1}} \cdots \sum_{x_n} p_{X_1,X_2,\ldots,X_n}(x_1, x_2, \ldots, x_n).$$

8: Conditional distributions

$$\begin{aligned} & p_{(X_1,\ldots,X_k)|(X_{k+1},\ldots,X_n)}(x_1, \ldots, x_k | x_{k+1}, \ldots, x_n) \\ &= \frac{p_{X_1,\ldots,X_n}(x_1, \ldots, x_n)}{p_{X_{k+1},\ldots,X_n}(x_{k+1}, \ldots, x_n)} \end{aligned}$$