LECTURE 6

TRANSFORMATION OF RANDOM VARIABLES

TRANSFORMATION OF FUNCTION OF A RANDOM VARIABLE

UNIVARIATE TRANSFORMATIONS

TRANSFORMATION OF RANDOM VARIABLES

- If X is an rv with cdf $F(x)$, then $Y=g(X)$ is also an rv.

- If we write $y=g(x)$, the function $g(x)$ defines a mapping from the original sample space of X, S, to a new sample space, $\mathcal{Y}$, the sample space of the rv Y.

$$g(x)\colon S \rightarrow \mathcal{Y}$$

1-TO-1 TRANSFORMATION OF RANDOM VARIABLES

- Let y=g(x) define a 1-to-1 transformation. That is, the equation y=g(x) can be solved uniquely:

$$x = g^{-1}(y)$$

- Ex: Y=X-1 → X=Y+1 1-to-1
- Ex: $Y=X^2$ → X=± sqrt(Y) not 1-to-1
- When transformation is not 1-to-1, find disjoint partitions of S for which transformation is 1-to-1.

If X is a discrete r.v. then S is countable. The sample space for Y=g(X) is $\mathcal{Y}$={y:y=g(x),x$\in$ S}, also countable. The pmf for Y is

$$f_Y(y) = P(Y = y) = \sum_{x \in g^{-1}(y)} P(X = x) = \sum_{x \in g^{-1}(y)} f(x)$$

Example

- Let X~GEO(p). That is, $f(x) = p(1-p)^{x-1} \text{ for } x = 1,2,3,...$
- Find the p.m.f. of Y=X-1
- Solution: X=Y+1

$$f_Y(y) = f_X(y+1) = p(1-p)^y \text{ for } y = 0,1,2,...$$

- P.m.f. of the number of failures before the first success
- Recall: X~GEO(p) is the p.m.f. of number of Bernoulli trials required to get the first success

Example

- Suppose X~Poisson with $p(x) = \frac{e^{-\lambda}\lambda^{x}}{x!} \quad x = 0,1,2,\ldots$

- Let Y=4X X=Y/4
- Then
$$p(x = y/4) = \frac{e^{-\lambda}\lambda^{y/4}}{(y/4)!} \quad y = 0,4,8,\ldots$$

CONTINUOUS RANDOM VARIABLE

- Let X be an rv of the continuous type with pdf f. Let $y=g(x)$ be differentiable for all x and non-zero. Then, $Y=g(X)$ is also an rv of the continuous type with pdf given by

$$h(y) = \begin{cases} f(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right| & for\ y \in \psi \\ 0 & o.w. \end{cases}$$

Example

- Let X have the density

$$f(x)=\begin{cases}1, & 0<x<1\\ 0, & \text{otherwise}\end{cases}$$

Let $Y=e^X$. $\mathrm{X}=\mathrm{g}^{-1}(\mathrm{y})=\ln\mathrm{Y}\rightarrow \mathrm{dx}=(1/\mathrm{y})\mathrm{dy}$.

$$h(y)=1.\left|\frac{1}{y}\right|, 0<\log y<1$$

$$h(y)=\begin{cases}\dfrac{1}{y}, & 1<y<e\\ 0, & \text{otherwise}\end{cases}$$

Example

- Suppose X has an exponential distribution with

$$f_X(x) = \frac{1}{\lambda} e^{-x/\lambda}, x > 0, \lambda > 0$$

- Let Y=4X-2 X=(Y+2)/4

$$\mathrm{X} = \mathrm{g}^{-1}(\mathrm{y}) = (\mathrm{Y}+2)/4 \rightarrow \mathrm{dx} = (1/4)\mathrm{dy}.$$

$$h(y) = \begin{cases} \frac{1}{\lambda} e^{-\lambda(y+2)/4} \left| \frac{1}{4} \right| y > -2, \lambda > 0 \\ 0 \qquad\qquad\qquad o.w. \end{cases}$$

TRANSFORMATION THAT ARE NOT 1-TO-1

- Let y=g(x) if the equation y=g(x) can not be solved uniquely then not one to one transformation should be used.

  When transformation is not 1-to-1, find disjoint partitions of *S* for which transformation is 1-to-1 and use 1-to-1 transformation or use cumulative method.

Example

- Let X be an rv with pmf

$$p(x)=\frac{4}{31}\left(\frac{1}{2}\right)^{x}\quad x=-2,-1,0,1,2$$

Let Y=|X| $\quad S=\{-2,-1,0,1,2\}\quad \mathcal{Y}=\{0,1,4\}$

$$p_y(0)=\frac{4}{31}$$

$$p_y(1)=p_x(-1)+p_x(1)=\frac{10}{31}$$

$$p_y(2)=p_x(-2)+p_x(2)=\frac{17}{31}$$

Example

- Let X be an rv with pmf

$$p(x) = \begin{cases} 1/5, x = -2 \\ 1/6, x = -1 \\ 1/5, x = 0 \\ 1/15, x = 1 \\ 11/30, x = 2 \end{cases}$$

Let Y=X^2. ⟶ $S = \{-2, -1, 0, 1, 2\}$ ⟶ $\mathcal{Y} = \{0, 1, 4\}$

$$p(y) = \begin{cases} 1/5, y = 0 \\ 7/30, y = 1 \\ 17/30, y = 4 \end{cases}$$

Example

- Stores located on a linear city with density f(x)=0.05 -10≤x ≤ 10, 0 otherwise
- Courier incurs a cost of $U=16X^2$ when she delivers to a store located at X (her office is located at 0)

$$U = u \quad \Rightarrow \quad 16X^2 = u \quad X = \pm\frac{\sqrt{u}}{4}$$

$$U \le u \quad \Rightarrow \quad -\frac{\sqrt{u}}{4} \le X \le \frac{\sqrt{u}}{4}$$

$$F_U(u) = P(U \le u) = \int_{-\sqrt{u}/4}^{\sqrt{u}/4} 0.05dx = 0.05\left(\frac{\sqrt{u}}{4} - \left(-\frac{\sqrt{u}}{4}\right)\right) = \frac{\sqrt{u}}{40} \quad 0 \le u \le 1600$$

$$f_U(u) = \frac{dF_U(u)}{du} = \frac{u^{-1/2}}{80} \quad 0 \le u \le 1600$$

Example

- Let X have the density

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, -\infty < x < \infty.$$

Let $W=X^2$. Find the pdf of W.

First step

Find the distribution function of W

$$G(w) = P[W \le w] = P[X^2 \le w]$$

$$= P\left[-\sqrt{w} \le X \le \sqrt{w}\right] \text{ if } w \ge 0$$

$$= \int_{-\sqrt{w}}^{\sqrt{w}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

$$= F\left(\sqrt{w}\right) - F\left(-\sqrt{w}\right)$$

where $$F'(x) = f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

Second step

Find the density function of W

$g(w) = G'(w)$.

$$= F'\left(\sqrt{w}\right)\frac{d\sqrt{w}}{dw} - F'\left(-\sqrt{w}\right)\frac{d\left(-\sqrt{w}\right)}{dw}$$

$$= f\left(\sqrt{w}\right)\frac{1}{2}w^{-\frac{1}{2}} + f\left(-\sqrt{w}\right)\frac{1}{2}w^{-\frac{1}{2}}$$

$$= \frac{1}{\sqrt{2\pi}}e^{-\frac{w}{2}}\frac{1}{2}w^{-\frac{1}{2}} + \frac{1}{\sqrt{2\pi}}e^{-\frac{w}{2}}\frac{1}{2}w^{-\frac{1}{2}}$$

$$= \frac{1}{\sqrt{2\pi}}w^{-\frac{1}{2}}e^{-\frac{w}{2}} \quad \text{if } w \geq 0.$$

TRANSFORMATION OF FUNCTION OF TWO OR MORE RANDOM VARIABLES

BIVARIATE TRANSFORMATIONS

DISCRETE CASE

- Let X_1 and X_2 be a bivariate random vector with a known probability distribution function.
- Consider a new bivariate random vector (U, V) defined by $U=g_1(X_1, X_2)$ and $V=g_2(X_1, X_2)$ where $g_1(X_1, X_2)$ and $g_2(X_1, X_2)$ are some functions of X_1 and X_2 .

- Then, the joint pmf of *(U,V)* is

$$f_{U,V}(u,v) = \Pr(U = u, V = v) = \sum_{(x_1,x_2) \in A_{U,V}} f_{X_1,X_2}(x_1, x_2)$$

EXAMPLE

- Let X_1 and X_2 be independent Poisson distribution random variables with parameters λ_1 and λ_2 with joint pmf .

$$p(x_1, x_2) = \frac{\lambda_1^{x_1} \lambda_2^{x_2} e^{-\lambda_1 - \lambda_2}}{x_1! x_2!} \quad x_1 = 0,1,2,..., \quad x_2 = 0,1,2,...,$$

- Find the distribution of $Y_1=X_1+X_2$. We need to define new variable $Y_2=X_2$. Then $Y_1=0,1,2,...$, $Y_2=0,1,2,...,Y_1$ and $X_2=Y_2$ and $X1=Y_1-Y_2$.

$$p(y_1, y_2) = \frac{\lambda_1^{y_1 - y_2} \lambda_2^{y_2} e^{-\lambda_1 - \lambda_2}}{(y_1 - y_2)! y_2!} \quad y_1 = 0,1,2,..., \quad y_2 = 0,1,2,..., y_1$$

$$p(y_1) = \sum_{y_2=0}^{y_1} p(y_1, y_2) = \frac{e^{-\lambda_1 - \lambda_2}}{y_1!} \sum_{y_2=0}^{y_1} \frac{\lambda_1^{y_1 - y_2} \lambda_2^{y_2} y_1!}{(y_1 - y_2)! y_2!} = \frac{(\lambda_1 + \lambda_2)^{y_1} e^{-\lambda_1 - \lambda_2}}{y_1!}$$

CONTINUOUS CASE

- Let $X=(X_1, X_2, \ldots, X_n)$ have a continuous joint distribution for which its joint pdf is f, and consider the joint pdf of new random variables $Y_1, Y_2, \ldots, Y_k$ defined as

$$\left.\begin{aligned} Y_1 &= g_1(X_1, X_2, \cdots, X_n) \\ Y_2 &= g_2(X_1, X_2, \cdots, X_n) \\ &\vdots \\ Y_k &= g_k(X_1, X_2, \cdots, X_n) \end{aligned}\right\} *$$

- If the transformation T is one-to-one and onto, then there is no problem of determining the inverse transformation, and we can invert the equation in (*) and obtain new equations as

$$\left.\begin{aligned} x_1 &= g_1^{-1}(y_1, y_2, \cdots, y_k) \\ x_2 &= g_2^{-1}(y_1, y_2, \cdots, y_k) \\ &\vdots \\ x_n &= g_n^{-1}(y_1, y_2, \cdots, y_{k=n}) \end{aligned}\right\} **$$

Assuming that the partial derivatives $\partial g_i^{-1} / \partial y_i$ exist at every point $(y_1, y_2, \ldots, y_{k=n})$.

- Under these assumptions, we have the following determinant *J*

$$J = det\begin{bmatrix} \frac{\partial g_1^{-1}}{\partial y_1} & \cdots & \frac{\partial g_1^{-1}}{\partial y_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_n^{-1}}{\partial y_1} & \cdots & \frac{\partial g_n^{-1}}{\partial y_n} \end{bmatrix}$$

called as the Jacobian of the transformation specified by (**). Then, the joint pdf of Y_1, Y_2,…,Y_k can be obtained by using the change of variable technique of multiple variables.

- As a result, the new p.d.f. is defined as follows:

$$g(y_1, y_2, \cdots, y_n) = \begin{cases} f_{X_1,\cdots,X_n}\left(g_1^{-1}, g_2^{-1}, \cdots, g_n^{-1}\right) | J |, \text{for} \left(y_1, y_2, \cdots, y_n\right) \in \Psi \\ 0, otherwise \end{cases}$$

Example

Let X1 and X2 ~ Exp(1)

$$f_{x_1,x_2}(x_1,x_2) = e^{-(x_1+x_2)} \quad x_1 > 0, x_2 > 0$$

Consider the random variables Y1=X1, Y2=X1+X2. Find pdf of Y1 and Y2?

X1=Y1 and X2=Y2-Y1 the Jacobian is $\begin{vmatrix} 1 & 0 \\ -1 & 1 \end{vmatrix} = 1$

Then

$$f_{y_1,y_2}(y_1,y_2) = f_{x_1,x_2}(y_1, y_2 - y_1) = e^{-y_2} \quad 0 < y_1 < y_2 < \infty$$

METHOD OF CONDITIONING

- $U=h(X_1,X_2)$
- Find $f(u|x_2)$ by transformations (Fixing $X_2=x_2$)
- Obtain the joint density of U, X_2:
  - $f(u,x_2) = f(u|x_2)f(x_2)$
- Obtain the marginal distribution of U by integrating joint density over X_2

$$f_U(u) = \int_{-\infty}^{\infty} f(u \mid x_2) f(x_2) dx_2$$

Example

- X_1~Beta($\alpha=2,\beta=2$) X_2~Beta($\alpha=3,\beta=1$) Independent
- $U=X_1X_2$
- Fix $X_2=x_2$ and get $f(u|x_2)$

$$f(x_1)=6x_1(1-x_1) \quad 0\le x_1\le 1 \quad f(x_2)=3x_2^2 \quad 0\le x_2\le 1$$

$$U=X_1x_2 \Rightarrow X_1=U/x_2 \quad \Rightarrow \quad \frac{dX_1}{dU}=1/x_2$$

$$f(u|x_2)=6(u/x_2)(1-u/x_2)\left|\frac{1}{x_2}\right| \quad 0\le u\le x_2$$

$$f(u,x_2)=f(u|x_2)f(x_2)=6(u/x_2)(1-u/x_2)\left|\frac{1}{x_2}\right|3x_2^2=18u\left(1-\frac{u}{x_2}\right) \quad 0\le u\le x_2\le 1$$

$$\Rightarrow f_U(u)=\int_u^1 f(u|x_2)f(x_2)dx_2=\int_u^1\left(18u-\frac{18u^2}{x_2}\right)dx_2=18ux_2-18u^2\ln(x_2)\Big|_u^1=(18u-0)-\left(18u^2-18u^2\ln(u)\right)$$

$$=18u(1-u+u\ln(u)) \quad 0<u\le 1$$

Example

- X_1, X_2 independent Exponential(θ)
- $f(x_i)=\theta^{-1}e^{-xi/\theta}$ $x_i>0$, $\theta>0$, $i=1,2$
- $f(x_1,x_2)= \theta^{-2}e^{-(x1+x2)/\theta}$ $x_1,x_2>0$
- $U=X_1+X_2$

$$U = u \Rightarrow X_1 + X_2 = u \Rightarrow X_1 = u - x_2$$

$$U \le u \Rightarrow X_1 + X_2 \le u \Rightarrow X_2 \le u, X_1 \le u - X_2$$

$$P(U \le u) = \int_0^u \int_0^{u-x_2} \frac{1}{\theta^2} e^{-x_1/\theta} e^{-x_2/\theta} dx_1 dx_2 = \int_0^u \frac{1}{\theta} e^{-x_2/\theta} \left(-e^{-x_1/\theta}\right)\Big|_0^{u-x_2} dx_2$$

$$= \int_0^u \frac{1}{\theta} e^{-x_2/\theta} \left[1 - e^{-(u-x_2)/\theta}\right] dx_2 = \int_0^u \frac{1}{\theta} e^{-x_2/\theta} dx_2 - \int_0^u \frac{1}{\theta} e^{-(x_2 - u + x_2)/\theta} dx_2$$

$$= \left(1 - e^{-u/\theta}\right) - \frac{1}{\theta} u e^{-u/\theta} \Rightarrow \quad f_U(u) = \frac{1}{\theta} e^{-u/\theta} - \left[\left(\frac{1}{\theta}\right) e^{-u/\theta} - \left(\frac{u}{\theta^2}\right) e^{-u/\theta}\right]$$

$$= \frac{1}{\theta^2} u e^{-u/\theta} \quad u > 0 \quad \Rightarrow \quad U \sim Gamma(\alpha = 2, \beta = \theta)$$

M.G.F. Method

- If $X_1, X_2, \ldots, X_n$ are independent random variables with MGFs $M_{X_i}(t)$, then the MGF of $Y = \sum_{i=1}^{n} X_i$ is $M_Y(t) = M_{X_1}(t) \ldots M_{X_n}(t)$

Example

Let $X_i \overset{independent}{\sim} Bin(n_i, p)$

Then find the pmf of $Y = \sum_{i=1}^{k} X_i$

$$M_Y(t) = M_{X_1}(t) \ldots M_{X_k}(t)$$

$$= (pe^t + q)^{n_1} \ldots (pe^t + q)^{n_k}$$

$$= (pe^t + q)^{n_1 + \ldots + n_k}$$

$$\sum_{i=1}^{k} X_i \sim Bin(n_1 + n_2 + \cdots + n_k, p).$$

$$M_Y(t) = M_{X_1}(t) \ldots M_{X_k}(t)$$

$$= (pe^t + q)^{n_1} \ldots (pe^t + q)^{n_k}$$

$$= (pe^t + q)^{n_1 + \ldots + n_k}$$

Example

$$X_i \sim Gamma(\alpha_i, \beta) \quad i = 1, \ldots, n \quad \text{(independent)}$$

$$M_{X_i}(t) = (1 - \beta t)^{-\alpha_i} \quad i = 1, \ldots, n$$

$$Y = \sum_{i=1}^{n} X_i$$

$$M_Y(t) = E\left(e^{tY}\right) = E\left(e^{t(X_1 + \ldots + X_n)}\right) = E\left(e^{tX_1} \cdots e^{tX_n}\right) = M_{X_1}(t) \cdots M_{X_n}(t)$$

$$(1 - \beta t)^{-\alpha_1} \cdots (1 - \beta t)^{-\alpha_n} = (1 - \beta t)^{-\sum_{i=1}^{n} \alpha_i}$$

$$\Rightarrow Y = \sum_{i=1}^{n} X_i \sim Gamma\left(\sum_{i=1}^{n} \alpha_i, \beta\right)$$

Example

$$X_i \sim Normal(\mu_i, \sigma_i^2) \quad i = 1,\ldots, n \quad \text{(independent)}$$

$$M_{X_i}(t) = \exp\left\{ \mu_i t + \frac{\sigma_i^2 t^2}{2} \right\} \quad i = 1,\ldots, n$$

$$Y = \sum_{i=1}^{n} a_i X_i \quad \{a_i\} \equiv \text{fixed constants}$$

$$M_Y(t) = E\left(e^{tY}\right) = E\left(e^{t(a_1 X_1 + \ldots + a_n X_n)}\right) = E\left(e^{ta_1 X_1} \cdots e^{ta_n X_n}\right) = M_{X_1}(a_1 t) \cdots M_{X_n}(a_n t)$$

$$\exp\left\{ \mu_1 a_1 t + \frac{\sigma_1^2 a_1 t^2}{2} \right\} \cdots \exp\left\{ \mu_n a_n t + \frac{\sigma_n^2 a_n t^2}{2} \right\} = \exp\left\{ \left(\sum_{i=1}^{n} a_i \mu_i\right) t + \frac{\left(\sum_{i=1}^{n} a_i^2 \sigma_i^2\right) t^2}{2} \right\}$$

$$\Rightarrow Y = \sum_{i=1}^{n} a_i X_i \sim Normal\left(\sum_{i=1}^{n} a_i \mu_i, \sum_{i=1}^{n} a_i^2 \sigma_i^2 \right)$$

ORDER STATISTICS

ORDER STATISTICS

- Let $X_1, X_2, \ldots, X_n$ be a r.s. of size n from a distribution of continuous type having pdf $f(x), a<x<b$. Let $X_{(1)}$ be the smallest of X_i, $X_{(2)}$ be the second smallest of $X_i, \ldots$, and $X_{(n)}$ be the largest of X_i.

$$a < X_{(1)} \le X_{(2)} \le \cdots \le X_{(n)} < b$$

  - $X_{(i)}$ is the i-th order statistic.

$$X_{(1)} = \min\{X_1, X_2, \cdots, X_n\}$$

$$X_{(n)} = \max\{X_1, X_2, \cdots, X_n\}$$

ORDER STATISTICS

- It is often useful to consider ordered random sample.
- Example: suppose a r.s. of five light bulbs is tested and the failure times are observed as (5,11,4,100,17). These will actually be observed in the order of (4,5,11,17,100). Interest might be on the k_{th} smallest ordered observation, e.g. stop the experiment after k_{th} failure. We might also be interested in joint distributions of two or more order statistics or functions of them.

JOINT PDF OF THE OREDER STATISTICS

- If $X_1, X_2, \ldots, X_n$ is a r.s. of size *n* from a population with *continuous* pdf *f(x),* then the joint pdf of the order statistics $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ is

$$g\left(x_{(1)}, x_{(2)}, \cdots, x_{(n)}\right) = n!\, f\left(x_{(1)}\right) f\left(x_{(2)}\right) \cdots f\left(x_{(n)}\right)$$

$$\text{for } x_{(1)} \leq \ldots \leq x_{(n)}$$

Order statistics are **not** independent.

The joint pdf of *ordered* sample is **not** same as the joint pdf of *unordered* sample.

Note: For discrete distributions, we need to take ties into account (two X’s being equal).

Example

Find the joint pdf of the order statistics for the uniform distribution, the standard exponential distribution and normal distribution?

Solution: p.d.f for the uniform is:

$$f(x) = 1, \quad 0 < x < 1$$

$$g(y_1, y_1, \ldots, y_n) = n!, \quad 0 < y_1 < y_1 < \ldots, < y_n < 1,$$

Solution: p.d.f for the standard Exponential distribution is:

$$f(x) = e^{-x}, \quad x > 0$$

$$g(y_1, y_1, \ldots, y_n) = n!e^{-\sum_{i=1}^{n} y_i}, \quad 0 < y_1 < y_1 < \ldots, < y_n < \infty,$$

Solution: p.d.f for the standard normal distribution is:

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad -\infty < x < \infty$$

$$g(y_1, y_1, \ldots, y_n) = \frac{n!}{\sqrt{2\pi}} e^{-\sum_{i=1}^{n} \frac{y_i^2}{2}}, \quad -\infty < y_1 < y_1 < \ldots, < y_n < \infty,$$

Example

- Suppose that X_1, X_2 and X_3 represent a random sample of size 3 from population with pdf

$$f(x) = 2x \quad 0 < x < 1$$

- Joint pdf of order statistics Y_1, Y_2 and Y_3?

$$g(y_1, y_2, y_3) = 3!(2y_1)(2y_2)(2y_3) = 48 y_1 y_2 y_3 \quad 0 < y_1 < y_2 < y_3 < 1$$

- Marginal pdf of Y_1

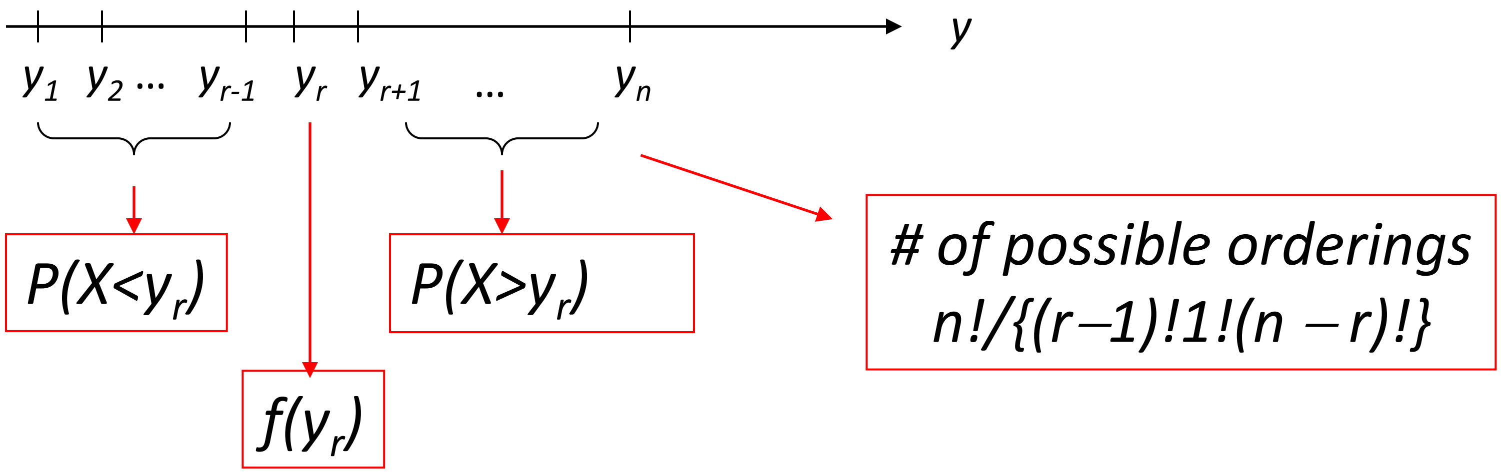
$$g_1(y_1) = \int_{y1}^{1} \int_{y2}^{1} 48 y_1 y_2 y_3 \, dy_2 dy_3 = 6 y_1 (1 - y_1{}^2)^2 \quad 0 < y_1 < 1$$

THE MARGINAL DISTRIBUTIONS FOR THE ODER STATISTICS

Theorem: (p.d.f of the rth order statistics) If X_1, $X_2,\ldots,X_n$ be a r.s. of size n from a population with *continuous* pdf $f(x)$, then the p.d.f. of the rth order statistics $X_{(r)}$ is given as:

$$g_r(y_r) = \frac{n!}{(r-1)!(n-r)!} f(y_r) \times [F(y_r)]^{r-1}[1-F(y_r)]^{n-1}, \ -\infty < y_r < \infty$$

- *r*-th Order Statistic



$$g_r(y_r) = \frac{n!}{(r-1)!(n-r)!} f(y_r)$$

$$\times [F(y_r)]^{r-1}[1-F(y_r)]^{n-1}, \quad -\infty < y_r < \infty$$

Example

- *X~Uniform(0,1)*. A r.s. of size *n* is taken. Find the p.d.f. of kth order statistic.
- Solution: Let Y_k be the kth order statistic.

$$g_{Y_k}(y) = \frac{n!}{(k-1)!(n-k)!} y^{k-1}(1-y)^{n-k}\ 1$$

$$= \frac{\Gamma(n+1)}{\Gamma(k)\Gamma(n-k+1)} y^{k-1}(1-y)^{n-k} \quad \text{for} \quad 0<y<1$$

$$Y_k \sim \text{Beta}(k, n-k+1)$$

Example

- Suppose that $X_1, X_2, \ldots, X_n$ represent a random sample of size n from population with pdf

$$f(x) = 2x \quad 0 < x < 1$$

- Marginal pdf of Y_1 and Y_n?

$$F(x) = x^2 \quad 0 < x < 1$$

$$g_1(y_1) = 2ny_1(1 - y_1{}^2)^{n-1} \; 0 < y_1 < 1$$
$$g_n(y_n) = 2ny_n(y_n{}^2)^{n-1} \; 0 < y_1 < 1$$

Theorem (p.d.f of the largest order statistics): If X_1, $X_2,\dots,X_n$ be a r.s. of size n from a population with *continuous* pdf $f(x)$, then the p.d.f. of the Largest order statistics $Y_{(n)}$ is given as:

$$g_n(y_n) = n f(y_n)[F(y_n)]^{n-1}, \quad -\infty < y_n < \infty$$

Theorem: (p.d.f of the smallest order statistics) If X_1, $X_2,\ldots,X_n$ be a r.s. of size n from a population with *continuous* pdf $f(x)$, then the p.d.f. of the smallest order statistics $X_{(1)}$ is given as:

$$g_1(y_1) = nf\,(y_1)[1-F(y_1)]^{n-1}, \quad -\infty < y_n < \infty$$

Example

Let $Y_1 < Y_2 < \ldots < Y_6$ are an O. S. of sample size *n = 6* and the p.d.f. of this sample is

$$f(x) = \frac{1}{2}, \quad 0 < x < 2$$

Find:

$$g_r(y_r) \quad, g_1(y_1) \;, g_6(y_6)$$

Solution:

$$g_r(y_r) = \frac{6!}{(r-1)!(6-r)!2^6} y_r^{r-1}[2-y_r]^{6-1}, \; 0 < y_r < 2$$

$$g_1(y_1) = \frac{6}{2^6}[2-y_1]^5, \; 0 < y_1 < 2$$

$$g_6(y_6) = \frac{6}{2^6} y_6^5, \; 0 < y_6 < 2$$

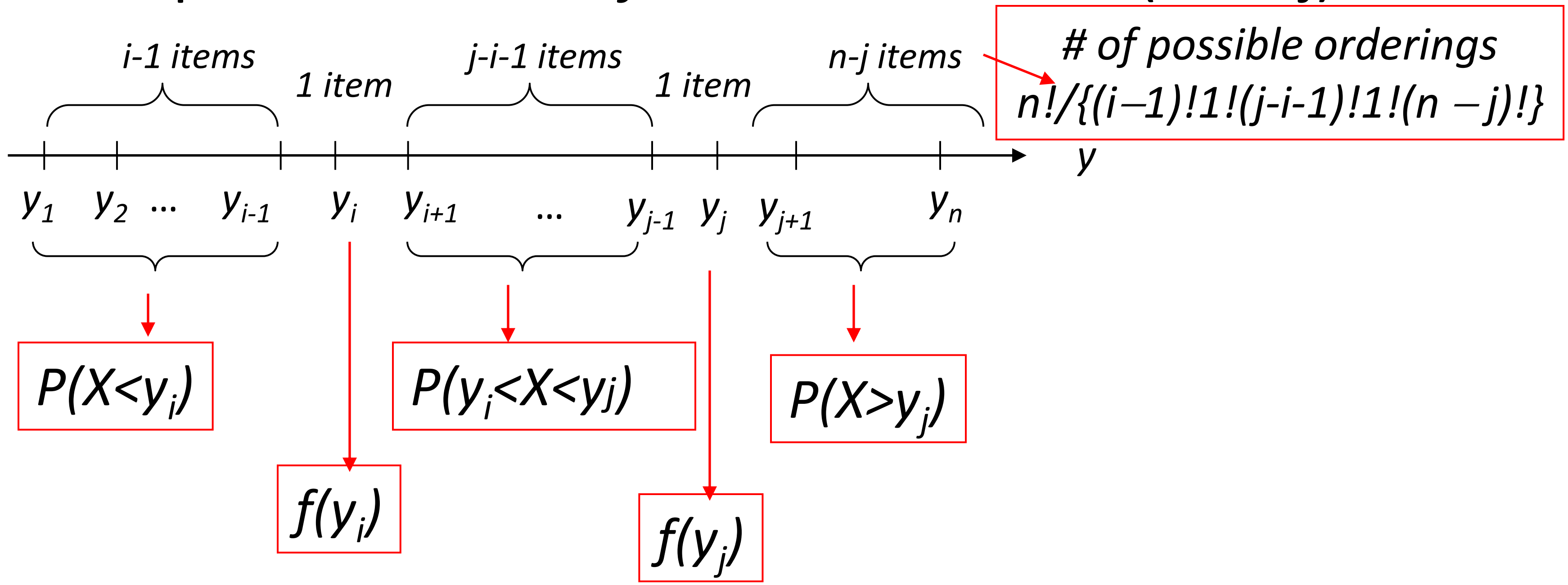
JOINT P.D.F. OF i-TH AND j-TH ORDER STATISTIC (FOR i<j)

Theorem:

If $X_1, X_2, \ldots, X_n$ be a r.s. of size *n* from a population with *continuous* pdf *f(x),* and $Y_1 < Y_2 < \ldots < Y_n$ are the order statistics of that sample, then the p.d.f. of the two order statistics $Y_i < Y_j$, i<j and i,j = 1,2, …,n is given as

$$g_{ij}(y_i, y_j) = \frac{n!}{(i-1)!(j-i-1)!(n-j)!}[F(y_i)]^{i-1} f(y_i)[F(y_j)-F(y_i)]^{j-i-1} f(y_j)[1-F(y_j)]^{n-j}$$

- Joint p.d.f. of *i*-th and *j*-th Order Statistic (for i<j)



$$g_{ij}(y_i, y_j) = \frac{n!}{(i-1)!(j-i-1)!(n-j)!}[F(y_i)]^{i-1} f(y_i)[F(y_j) - F(y_i)]^{j-i-1} f(y_j)[1-F(y_j)]^{n-j}$$

Example

- Suppose that $X_1, X_2, \ldots, X_n$ represent a random sample of size n from population with pdf

$$f(x) = 2x \quad 0 < x < 1$$

- Find the density of range $R = Y_n - Y_1$?

$$F(x) = x^2 \quad 0 < x < 1$$

$$g_{1,n}(y_1, y_n) = \frac{n!}{(n-2)!} 2y_1 2y_n ({y_n}^2 - {y_1}^2)^{n-2} \qquad 0 < y_1 < 1$$