Section :- Function of Random sample / Random vector.

Set up :- Let $X_1, \ldots, X_n$ be an r.s. from the distⁿ. F, (f: joint density f^n.).

OR $(X_1, \ldots, X_n) \xrightarrow{\text{Joint density}} f_{X_1, \ldots, X_n}(x_1, \ldots, x_n)$

Q :- Suppose, we want to know the CDF (or PDF) of $U(X_1, \ldots, X_n)$.

Three approaches :-

1. CDF based approach.
2. Jacobian based approach.
3. MGF approach.

__Approach 1:-__ (CDF based approach) :-

Let F_U be the CDF of $U(X_1, \ldots, X_n)$.

$$F_U(y) = P\left[\omega : U\left(X_1^{(\omega)}, \ldots, X_n^{(\omega)}\right) \leq y\right]$$

$$= \underset{\{(x_1, \ldots, x_n) :\ U(x_1, \ldots, x_n) \leq y\}}{\int\int \cdots \int} f_{X_1, X_2, \ldots, X_n}(x_1, \ldots, x_n)\, dx_1\, dx_2 \cdots dx_n$$

Hence, $f_U(y) = \frac{d}{dy} F_U(u)$.

f_U: P.D.F. of U

__Example 1:-__ $X_1, \ldots, X_n$ indep. Poisson (λ).

Want to know the distn. / p.m.f. of $\sum_{i=1}^{n} X_i$.

Consider $n=2$. $\quad Y = X_1 + X_2, \quad X_1 \perp\!\!\!\perp X_2,$

$X_1 \sim P(\lambda_1)$

$X_2 \sim P(\lambda_2)$.

$$P[Y=y] = P[X_1+X_2=y] \; ; \quad y = 0, \ldots$$

$$= \sum_{x=0}^{y} P[X_1 = x, X_2 = y-x]$$

$$= \sum_{x=0}^{y} P[X_1 = x]\, P[X_2 = y-x] \quad \text{since } X_1 \perp\!\!\!\perp X_2.$$

$$= \sum_{x=0}^{y} \frac{e^{-\lambda_1} \lambda_1^{x}}{x!} \times \frac{e^{-\lambda_2} \lambda_2^{y-x}}{(y-x)!}$$

$$= e^{-(\lambda_1+\lambda_2)} \frac{1}{y!} \sum_{x=0}^{y} \frac{y!}{x!\,(y-x)!} \lambda_1^{x} \lambda_2^{y-x}$$

$$= \frac{e^{-(\lambda_1+\lambda_2)}}{y!} \sum_{x=0}^{y} \binom{y}{x} \lambda_1^{x} \lambda_2^{y-x} = \frac{e^{-(\lambda_1+\lambda_2)}}{y!} \times (\lambda_1+\lambda_2)^{x+(y-x)}$$

$$= \frac{e^{-(\lambda_1+\lambda_2)} \times (\lambda_1+\lambda_2)^{y}}{y!} \Rightarrow Y \sim \text{Poi}(\lambda_1+\lambda_2).$$

Remark:- This result can be extended for any finite dimensional case (as long as the components are indep).

A few examples of f^{n} of random samples:-

R.S:- $(X_1, \ldots, X_n)$.

i) Sample mean $= \frac{1}{n}\sum_{i=1}^{n} X_i$

ii) Sample variance $= \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \frac{1}{n}\sum_{i=1}^{n} X_i\right)^2$.

iii) Maximum order statistic :- $X_{(n)} = \operatorname{Max}\{X_1, \ldots, X_n\}$.

iv) Minimum " " :- $X_{(1)} = \operatorname{Min}\{X_1, \ldots, X_n\}$.

v) i-th " " :- $X_{(i)} =$ i-th minimum among $\{X_1, \ldots, X_n\}$.

vi) Median $= X_{\left(\frac{n+1}{2}\right)}$ if n is odd

$$= \frac{X_{\left(\frac{n}{2}\right)} + X_{\left(\frac{n}{2}+1\right)}}{2} \quad \text{if } n \text{ is even.}$$

vii) Range $= X_{(n)} - X_{(1)}$

viii) Inter quartile range $= X_{\left(\left[\frac{3n}{4}\right]\right)} - X_{\left(\left[\frac{n}{4}\right]\right)}$

box f^n.

Examples :- Derivation of the CDF/PDF of $X_{(n)}$ & $X_{(1)}$.

Set up :- $X_1, \cdots, X_n \sim F$ and its p.d.f. is f.

Want to know the ~~d~~ p.d.f. of

$$Y = X_{(n)} = \text{Max}\{X_1, \cdots, X_n\}.$$

$$\& \; Z = X_{(1)} = \text{Min}\{X_1, \cdots, X_n\}.$$

i) $F_Y(y) = P[\omega : Y(\omega) \le y]$

CDF of Y

$= P[\text{Max}\{X_1, \ldots, X_n\} \le y]$.

$= P[X_1 \le y, X_2 \le y, \ldots, X_n \le y]$.

$= \prod_{i=1}^{n} P[X_i \le y]$ (since X_i's are indep).

$= \prod_{i=1}^{n} F_X(y) = \{F_X(y)\}^n$.

F_X → CDF of X, when $X_1, \ldots, X_n \overset{d}{=} X$.

$f_Y(y) = \frac{d}{dy} F_Y(y)$

p.d.f. of Y

$= \frac{d}{dy}\{F_X(y)\}^n = n\{F_X(y)\}^{n-1} f_X(y)$

f_X: p.d.f. of X.

ii) Now, we want to derive the density fn. Z.

$$F_Z(z) = P[\omega : Z(\omega) \leq z]$$

$$= P[\mathrm{Min}\{X_1, \ldots, X_n\} \leq z].$$

$$= 1 - P[\mathrm{Min}\{X_1, \ldots, X_n\} \geq z].$$

$$= 1 - P[X_1 \geq z, X_2 \geq z, \ldots, X_n \geq z].$$

$$= 1 - \prod_{i=1}^{n} P[X_i \geq z] \quad \text{since } X_i\text{'s are indep.}$$

$$= 1 - \prod_{i=1}^{n} \{1 - P[X_i \leq z]\}.$$

$$= 1 - \prod_{i=1}^{n} \{1 - F_X(z)\} = 1 - \{1 - F_X(z)\}^n.$$

Hence, $f_Z(z) = \frac{d}{dz} F_Z(z) = n\{1 - F_X(z)\}^{n-1} f_X(z).$

p.d.f. of Z

ii) Jacobian based approach:-

X_1, X_2
$Y_1 = X_1 - X_2$
$Y_2 = X_1 + X_2$

Set up:- $(X_1, \dots, X_n)$: n R.Vs.

Joint p.d.f. $\longrightarrow$ $f_{X_1,\dots,X_n}(x_1, \dots, x_n)$.

Want to know the form of joint density of

$Y_i = Y_i(X_1, \dots, X_n)$.

Solution:- Let the inverse transformation exists

i.e, $X_i = X_i(Y_1, \dots, Y_n)$.

Assumption: $\frac{\partial x_i}{\partial y_j}$ exists for i, j.

ii) $$J = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \cdots & \frac{\partial x_1}{\partial y_n} \\ \vdots & & \\ \frac{\partial x_n}{\partial y_1} & \cdots & \frac{\partial x_n}{\partial y_n} \end{vmatrix} \neq 0.$$

Now, the new random variables are

$$\left.\begin{aligned} Y_1 &= y_1(X_1, \dots, X_n) \\ &\vdots \\ Y_n &= y_n(X_1, \dots, X_n) \end{aligned}\right\} \text{ inverse } \quad \begin{aligned} X_1 &= x_1(Y_1, \dots, Y_n) \\ &\vdots \\ X_n &= x_n(Y_1, \dots, Y_n). \end{aligned}$$

Finally, the joint density of $(Y_1, \dots, Y_n)$ will be

$$f_{Y_1, \dots, Y_n}(y_1, \dots, y_n) = f_{X_1, \dots, X_n}\Big(x_1(y_1, \dots, y_n), \dots, x_n(y_1, \dots, y_n)\Big) \times |J|.$$

Example:- Let X_1, X_2 i.i.d. with p.d.f $\underset{X_1 = X_2 = X}{f_X}(x) = e^{-x}, \; x \geq 0$.

Question:- Want to know the density f^n of $X_1 - X_2$.

Consider the transformation :-

$$U = X_1 + X_2, \quad V = X_1 - X_2 \implies X_1 = \frac{U+V}{2} \quad \& \quad X_2 = \frac{U-V}{2}.$$

$$J = \begin{vmatrix} \frac{\partial x_1}{\partial u} & \frac{\partial x_1}{\partial v} \\ \frac{\partial x_2}{\partial u} & \frac{\partial x_2}{\partial v} \end{vmatrix} = \begin{vmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{vmatrix} = -\frac{1}{2}.$$

joint p.d.f. (X_1, X_2)

Next, $f_{X_1, X_2}(x_1, x_2) = e^{-x_1 - x_2}$, $x_1 \geq 0$, $x_2 \geq 0$.

This implies, the joint density of (U, V) is

joint p.d.f. of (U, V) $f_{U,V}(u,v) = e^{-u}|J| = e^{-u} \times \frac{1}{2} = \frac{1}{2}e^{-u}$. for range in (u,v).