Mathematical Statistics Lecture Notes — Chapter: 4.3 & 4.6

Transformations of Continuous Random Variables

Remember, we saw "Distribution Function Technique" to transform random variables. We'll see a new technique, whose logic is similar, but which is more practical.

The **change of variable** technique is as follows.

For a single Random Variable, (Marginal Distribution); let we have a Continuous Random Variable $X \in \mathcal{A}$ with pdf $f(x)$.

Also let, $Y = v(X)$. pdf of Y is found as follows;

(i) Find $x = w(y) = v^{-1}(y)$ and identify $y \in \mathcal{B}$

(ii) Find $|J| = |w'(y)|$ where J is the Jacobian of the transformation.

(iii)
$$\boxed{g(y) = \begin{cases} f[w(y)]\,|J| & y \in \mathcal{B} \\ 0 & \text{o.w.} \end{cases}}$$

4.26. If the p.d.f. of X is $f(x) = 2xe^{-x^2}$, $0 < x < \infty$, zero elsewhere, determine the p.d.f. of $Y = X^2$.

4.26)
$$f(x) = \begin{cases} 2x \cdot e^{-x^2} & 0 < x < \infty \\ 0 & \text{o.w.} \end{cases}$$

$y = x^2 \Rightarrow g(y) = ?$

$x = \sqrt{y} = w(y) \quad 0 < y < \infty$

$w'(y) = J = \dfrac{1}{2\sqrt{y}}$

$g(y) = f[w(y)] \cdot |J|, \; y > 0$

$g(y) = 2\sqrt{y} \cdot e^{-y} \cdot \dfrac{1}{2\sqrt{y}}, \; y > 0$

$g(y) = e^{-y}, \; y > 0$

Alp Giray Özen | 0533 549 91 08 | alp@lecturemania.com | www.lecturemania.com

Now, we extend the idea for more than one variables.

Let X_1, X_2 have the joint distribution $f(x_1, x_2)$, $(x_1, x_2) \in A$

also let $y_1 = u_1(X_1, X_2)$ and $y_2 = u_2(X_1, X_2)$

The joint distribution of y_1, y_2 is found as follows;

(i) Find $x_1 = w_1(y_1, y_2)$; $x_2 = w_2(y_1, y_2)$ and identify $(y_1, y_2) \in B$

(ii) Find

$$J = \begin{array}{c c} & \begin{array}{cc} y_1 & y_2 \end{array} \\ \begin{array}{c} x_1 \\ x_2 \end{array} & \begin{vmatrix} \dfrac{\partial x_1}{\partial y_1} & \dfrac{\partial x_1}{\partial y_2} \\ \dfrac{\partial x_2}{\partial y_1} & \dfrac{\partial x_2}{\partial y_2} \end{vmatrix} \end{array}$$

(iii)
$$\boxed{g(y_1, y_2) \begin{cases} = f[w_1(y_1, y_2), w_2(y_1, y_2)] \cdot |J| & (y_1, y_2) \in B \\ = 0 & \text{o.w.} \end{cases}}$$

If we want to find marginal distributions of y_1 and y_2, we follow

$$\boxed{g_1(y_1) = \int_{(y_2)} g(y_1, y_2)\, dy_2} \text{ and } \boxed{g_2(y_2) = \int_{(y_1)} g(y_1, y_2)\, dy_1}$$

Here we should point out a corollary.

Remember, if A and B are independent events,

$$P(A \cap B) = P(A) \cdot P(B)$$

Likewise, if X_1 and X_2 are independent random variables; $f(x_1, x_2) = f_1(x_1) \cdot f_2(x_2)$. It follows that, if $X_1, X_2, \ldots, X_n$ are a random sample from X with pdf $f(x)$, then $f(x_1, x_2, \ldots, x_n) = f(x_1) \cdot f(x_2) \cdot \ldots \cdot f(x_n)$

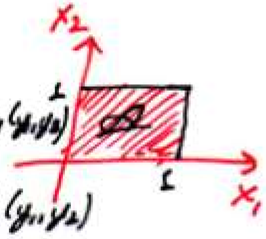
Alp Giray Özen | 0533 549 91 08 | alp@lecturemania.com | www.lecturemania.com

Example: let X has pdf $f(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & o.w. \end{cases}$

(Namely, $X \sim Uniform(0;1)$). A random sample of size 2 is drawn from this distribution. Find the joint pdf of $Y_1 = X_1 + X_2$ and $Y_2 = X_1 - X_2$. Also find marginal distributions of Y_1 and Y_2.

Answer: $h(x_1, x_2) = f(x_1) \cdot f(x_2) = 1 \qquad \overbrace{0 < x_1 < 1 \,;\, 0 < x_2 < 1}^{A}$

$\qquad = 0 \qquad o.w.$

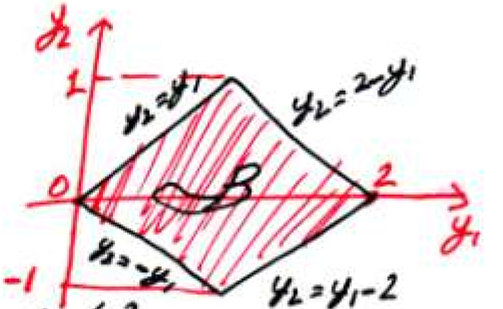


we have; $\left.\begin{aligned} y_1 &= x_1 - x_2 \\ y_2 &= x_1 + x_2 \end{aligned}\right\} \Rightarrow \begin{aligned} x_1 &= \tfrac{1}{2}(y_1 + y_2) = w_1(y_1, y_2) \\ x_2 &= \tfrac{1}{2}(y_1 - y_2) = w_2(y_1, y_2) \end{aligned}$

$0 < x_1 < 1$ and $0 < x_2 < 1$

$0 < \frac{1}{2}(y_1 + y_2) < 1 \qquad 0 < \frac{1}{2}(y_1 - y_2) < 1$

$0 < y_1 + y_2 < 2 \qquad 0 < y_1 - y_2 < 2$



$0 < y_1 + y_2 \qquad y_1 + y_2 < 2 \qquad 0 < y_1 - y_2 \qquad y_1 - y_2 < 2$

$\boxed{-y_1 < y_2} \qquad \boxed{y_2 < 2 - y_1} \qquad \boxed{y_2 < y_1} \qquad \boxed{y_1 - 2 < y_2}$

$$J = \begin{array}{c|cc|} & y_1 & y_2 \\ x_1 & 1/2 & 1/2 \\ x_2 & 1/2 & -1/2 \end{array} = -\frac{1}{4} - \frac{1}{4} = -\frac{1}{2} \Rightarrow |J| = \frac{1}{2}$$

$g(y_1, y_2) = h\left[w_1(y_1, y_2), w_2(y_1, y_2)\right] \cdot |J|$

$g(y_1, y_2) = h\left[\frac{1}{2}(y_1 + y_2), \frac{1}{2}(y_1, y_2)\right] \cdot |J|$

$g(y_1, y_2) = 1 \cdot \frac{1}{2} = \frac{1}{2} \quad (y_1, y_2) \in B$

$$g_1(y_1) = \begin{cases} \int_{-y_1}^{y_1} \frac{1}{2}\, dy_2 = y_1 & 0<y_1\le 1 \\ \int_{y_1-2}^{2-y_2} \frac{1}{2}\, dy_2 = 2-y_1 & 1<y_1<2 \end{cases}$$

$$g_1(y_1) = \begin{cases} y_1 & 0<y_1\le 1 \\ 2-y_1 & 1<y_1<2 \end{cases}$$

$$g_2(y_2) = \begin{cases} \int_{-y_2}^{y_2+2} \frac{1}{2}\, dy_1 = y_2+1 & -1<y_2\le 0 \\ \int_{y_2}^{2-y_2} \frac{1}{2}\, dy_1 = 1-y_2 & 0<y_2<1 \end{cases}$$

$$g_2(y_2) = \begin{cases} y_2+1 & -1<y_2\le 0 \\ 1-y_2 & 0<y_2<1 \end{cases}$$

4.33. Let X_1 and X_2 have the joint p.d.f. $h(x_1, x_2) = 2e^{-x_1-x_2}$, $0 < x_1 < x_2 < \infty$, zero elsewhere. Find the joint p.d.f. of $Y_1 = 2X_1$ and $Y_2 = X_2 - X_1$ and argue that Y_1 and Y_2 are independent.

4.33)

$$h(x_1,x_2) \begin{cases} = 2e^{-x_1-x_2} & 0<x_1<x_2 \\ = 0 & \text{o.w.} \end{cases}$$

$$\left.\begin{aligned} y_2 &= 2X_1 \\ y_2 &= X_2 - X_1 \end{aligned}\right\} \Rightarrow \begin{aligned} X_1 &= w_1(y_1,y_2) = \frac{y_1}{2} \\ X_2 &= w_2(y_1,y_2) = \frac{y_1}{2} + y_2 \end{aligned}$$

$$J = \begin{array}{c|cc|} & y_1 & y_2 \\ x_1 & 1/2 & 0 \\ x_2 & 1/2 & 1 \end{array} = \frac{1}{2}$$

$0 < x_1 < x_2$

$0 < y_1/2 < y_2$

$$\underbrace{0 < y_1}_{y_1>0} < \underbrace{y_1 + 2y_2}_{y_2>0}$$

$$g(y_1,y_2) = h\left(y_1/2,\ y_1/2 + y_2\right)\cdot\frac{1}{2} = 2\cdot e^{-\frac{y_1}{2}-\frac{y_1}{2}-y_2}\cdot\frac{1}{2}$$

$$g(y_1,y_2) = e^{-y_1-y_2} \quad \text{and} \quad g(y_1,y_2) = \underbrace{e^{-y_1}}_{\sim Exp(1)}\cdot \underbrace{e^{-y_2}}_{\sim Exp(1)}$$

So, Y_1 and Y_2 are independent

Alp Giray Özen | 0533 549 91 08 | alp@lecturemania.com | www.lecturemania.com

Order Statistics

Let $X_1, X_2, \ldots, X_n$ is a random sample from the distribution of the continuous type: $f(x)$. $a < x < b$

Also let, we ordered the sample values ~~from~~ in ascending order so that $y_1 = Min(X_1, X_2, \ldots, X_n)$; $y_n = Max(X_1, X_2, \ldots, X_n)$ and so y_k is the k^{th} rank order.

The joint pdf of $y_1, y_2, \ldots, y_k$ is given by,

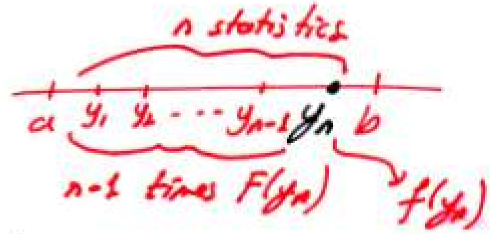
(I) $$\boxed{g(y_1, y_2, \ldots, y_n) = n!\,.\,f(y_1)\,f(y_2)\ldots\ldots f(y_n)} \quad a < y_1 < y_2 < \ldots < y_n < b$$

The idea is simple, n values can be ordeed in $n!$ different ways.

Next, consider the marginal distributions of y_k, $k = 1, 2, \ldots, n$. Keep in mind that, for statistics less than k^{th} order, we'll use $F(y_k) = P(Y_k \leq y_k)$ and for statistics more than k^{th} order, $1 - F(y_k) = P(Y_k \geq y_k)$.

Then;

(i) $$g_n(y_n) = n!\,.\,\frac{[F(y_n)]^{n-1}}{(n-1)!}\,.\,f(y_n) \qquad a < y_n < b$$

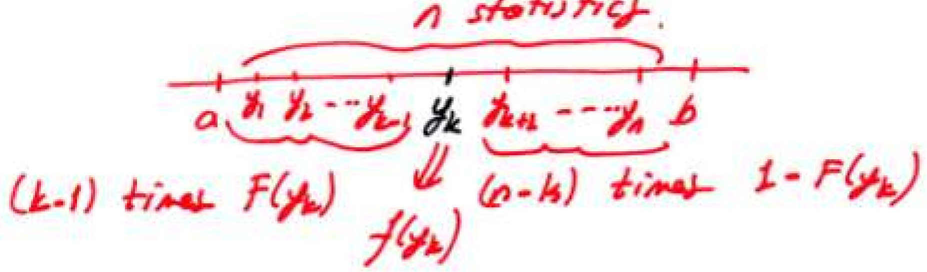


$$g_n(y_n) = n\,.\,[F(y_n)]^{n-1}\,.\,f(y_n) \qquad a < y_n < b$$
$$= 0 \qquad \text{o.w.}$$

(ii) $$g_1(y_1) = n!\,.\,\frac{[1 - F(y_1)]^{n-1}}{(n-1)!}\,f(y_1) = n\,.\,[1 - F(y_1)]^{n-1}\,.\,f(y_1) \qquad a < y_1 < b$$
$$= 0 \qquad \text{o.w.}$$

(iii) $$g_k(y_k) = \frac{n!}{(k-1)!\,.\,(n-k)!}\,.\,[F(y_k)]^{k-1}\,.\,[1-F(y_k)]^{n-k}\,.\,f(y_k) \qquad a<y_k<b$$

$$= 0 \qquad \text{o.w.}$$

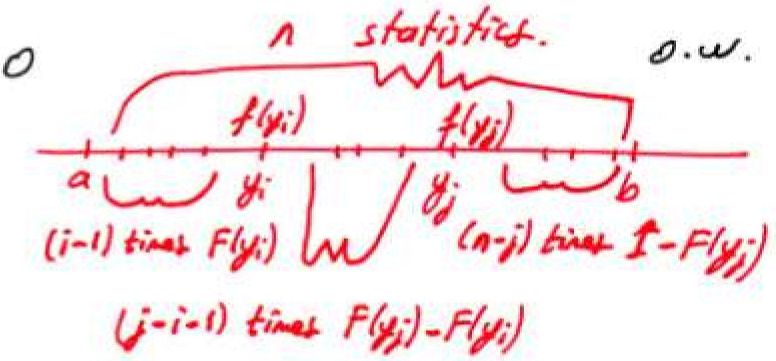


(II) Finally, the joint pdf of Y_i and Y_j is; ($y_i < y_j$)

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

$$a<y_i<y_j<b$$

$$= 0 \qquad \text{o.w.}$$



4.56. Let $Y_1 < Y_2 < Y_3 < Y_4$ be the order statistics of a random sample of size 4 from the distribution having p.d.f. $f(x) = e^{-x}$, $0 < x < \infty$, zero elsewhere. Find Pr $(3 \le Y_4)$.

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

$$F(x) = \int_0^x f(w)\,dw = \int_0^x e^{-w}\,dw = -e^{-w}\Big]_0^x = 1-e^{-x} \quad x>0$$

$$g_4(y_4) = 4\,[F(y_4)]^{n-1} \cdot f(y_4) = 4 \cdot [1-e^{-y_4}]^3 \cdot e^{-y_4} \quad y_4 > 0$$

$$P(Y_4 \le 3) = \int_0^3 g_4(y_4)\,dy_4 = \int_0^3 4 \cdot (1-e^{-y_4})^3 \cdot e^{-y_4}\,dy_4$$

$u = 1-e^{-y_4}$
$du = +e^{-y_4}\,dy_4$

$$= \int_0^{1-e^{-3}} 4u^3\,du = \frac{u^4}{4}\Big]_0^{1-e^{-3}} = \frac{1}{4} \cdot (1-e^{-3})^4 = 0,204$$

4.60. Let $Y_1 < Y_2 < \cdots < Y_n$ be the order statistics of a random sample of size n from a distribution with p.d.f. $f(x) = 1$, $0 < x < 1$, zero elsewhere. Show that the kth order statistic Y_k has a beta p.d.f. with parameters $\alpha = k$ and $\beta = n - k + 1$.

4.60) Beta distribution is defined as;

$$X \sim Beta(\alpha;\beta)$$

$$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\cdot\Gamma(\beta)} \cdot x^{\alpha-1} \cdot (1-x)^{\beta-1} \quad 0<x<1$$

$= 0 \quad$ o.w

where $\Gamma(w) = \int_0^\infty x^{w-1} \cdot e^{-x}\,dx$ and $\Gamma(w) = (w-1)!$
when w is natural #.

We have,

$$f(x) = \begin{cases} 1 & 0<x<1 \\ 0 & o.w \end{cases} \quad \text{and} \quad F(x) = \int_0^x \underbrace{f(w)}_{=1}dw = w\Big]_0^x = x \quad 0<1<x$$

Then, $$g_k(y_k) = \frac{n!}{(k-1)!\cdot(n-k)!} \cdot x^{k-1} \cdot (1-x)^{n-k} \cdot 1, \quad 0<y_k<1$$

$$g_k(y_k) = \frac{\Gamma(n+1)}{\Gamma(k)\cdot\Gamma(n-k+1)} \cdot x^{k-1} \cdot (1-x)^{n-k+1-1} \quad 0<y_k<1$$

So; $Y_k \sim Beta(\alpha = k;\ \beta = n-k+1)$

4.62. Find the probability that the range of a random sample of size 4 from the uniform distribution having the p.d.f. $f(x) = 1$, $0 < x < 1$, zero elsewhere, is less than $\frac{1}{2}$.

4.62) $f(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{o.w} \end{cases}$ and $F(x) = x \quad 0 < x < 1$

$$g_{1,4}(y_1, y_4) = \frac{4!}{0!.\,2!.\,0!}.\,[F(y_4) - F(y_1)]^2.\, f(y_4).\,f(y_1)$$

$$g_{1,4}(y_1, y_4) = 12.\,(y_4 - y_1)^2.\,1.\,1 = 12.\,(y_4 - y_1)^2 \quad 0 < y_1 < y_4 < 1$$

Range is $R = y_4 - y_1$

We also let $S = y_1$ (to construct a joint distribution)

Then, The joint distribution of R and S is,

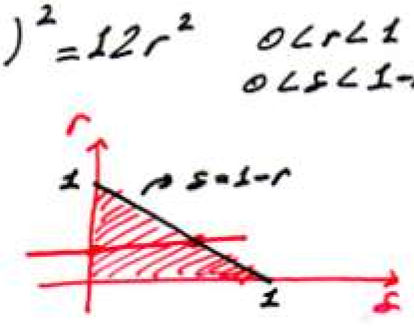
$$h(r,s) = g_{1,4}\,[w_1(r,s),\, w_2(r,s)].\,|J|$$

$$\left.\begin{aligned} R &= y_4 - y_1 \\ S &= y_1 \end{aligned}\right\} \Rightarrow \begin{aligned} y_1 &= w_1(R,S) = S \\ y_4 &= w_2(R,S) = R + S \end{aligned} \qquad J = \begin{array}{c|cc} & y_1 & y_4 \\ \hline r & -1 & 1 \\ s & 1 & 0 \end{array} = -1$$

$|J| = 1$

$$h(r,s) = g_{1,4}\,[s,\, r+s].\,1 = 12.\,(r - s + s)^2 = 12r^2 \quad \begin{array}{l} 0 < r < 1 \\ 0 < s < 1-r \end{array}$$

$$h_r(r) = \int_0^{1-r} h(r,s)\,ds = \int_0^{1-r} 12r^2\,ds = 12r^2 s\Big]_0^{1-r}$$



$$h_r(r) = 12r^2(1-r) \quad 0 < r < 1$$

Finally, $P(R < \frac{1}{2}) = \int_0^{1/2} 12r^2(1-r)\,dr = \left(\frac{12r^3}{3} - \frac{12r^4}{4}\right]_0^{1/2} = 12.\left(\frac{0{,}5^3}{3} - \frac{0{,}5^4}{4}\right)$

$= 0{,}3125$

4.64. If a random sample of size 2 is taken from a distribution having p.d.f. $f(x) = 2(1 - x)$, $0 < x < 1$, zero elsewhere, compute the probability that one sample observation is at least twice as large as the other.

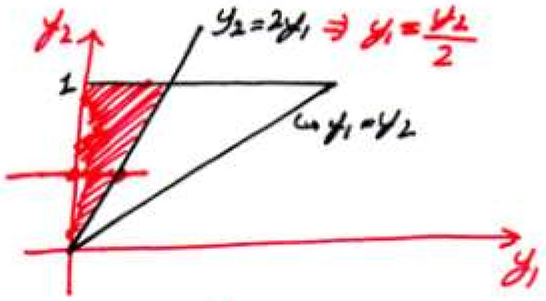
4.64) $f(x) = \begin{cases} 2(1-x) & 0<x<1 \\ 0 & \text{o.w.} \end{cases}$ and $F(x) = \int_0^x (2-2w)\,dw$

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

$g(y_1,y_2) = 2!\,.\,f(y_1)\,.\,f(y_2) = 2\,.\,2(1-y_1)\,.\,2(1-y_2)$

$= 8(1-y_1)(1-y_2) \quad 0<y_1<y_2<1$

$P(y_2 > 2y_1) = \int_0^1 \int_0^{y_2/2} 8(1-y_1)(1-y_2)\,dy_1\,dy_2$



$= 8\int_0^1 (1-y_2) \left[\underbrace{\int_0^{y_2/2} (1-y_1)\,dy_1}_{} \right] dy_2$

$= y_1 - \frac{y_1^2}{2} \Big]_0^{y_2/2} = \frac{y_2}{2} - \frac{y_2^2}{4} = \frac{1}{4}(2y_2 - y_2^2)$

$= 2\int_0^1 (1-y_2)(2y_2 - y_2^2)\,dy_2 = 2\int_0^1 (2y_2 - 3y_2^2 + y_2^3)\,dy_2 = 2\left(y_2^2 - y_2^3 + \frac{y_2^4}{4}\right]_0^1 = \frac{1}{2}$

4.70. Let X and Y denote independent random variables with respective probability density functions $f(x) = 2x$, $0 < x < 1$, zero elsewhere, and $g(y) = 3y^2$, $0 < y < 1$, zero elsewhere. Let $U = \min(X, Y)$ and $V = \max(X, Y)$. Find the joint p.d.f. of U and V.

Hint: Here the two inverse transformations are given by $x = u$, $y = v$ and $x = v$, $y = u$.

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

$g(y) = 3y^2 \quad 0<y<1 \qquad G(y) = y^3 \quad 0<y<1$

$U = \min(X, Y) \qquad V = \max(X, Y)$

let pdf : $m(u)$ $\qquad$ let pdf : $h(v)$

$$H(v)=P(\overset{max(X,Y)}{\widetilde{V}}\le v)=P(X\le v,\ Y\le v)\overset{\text{by independence}}{=}P(X\le v).P(Y\le v)$$

$$=F(v).G(v)=v^2.v^3=v^5 \quad 0<v<1$$

$V \overset{\shortparallel}{=} max(x,y)$ then, $h(v)=\frac{d}{dv}H(v)=5v^4 \quad 0<v<1$

$$M(u)=P(U\le u)=1-P(\overset{min(X,Y)}{\widetilde{U}}\ge u)=1-P(X\ge u,\ Y\ge u)$$

$$\overset{\text{by independence}}{=}1-P(X\ge u).P(Y\ge u)=1-[(1-P(X\le u)).(1-P(Y\le u))]$$

$$=1-[(1-F(u)).(1-G(u))]=1-[(1-u^2).(1-u^3)]$$

$$=u^2+u^3-u^5 \quad 0<u<1$$

$U \overset{\shortparallel}{=} min(x,y)$ then, $m(u)=\frac{d}{du}M(u)=2u+3u^2-5u^4 \quad 0<u<1$