Math 481/581 Mathematical Statistics I

Notes 6.

(p.1) <u>Transformation of a random pair (X, Y).</u>

Given a pair of random variables (X, Y) with joint density $f(x,y)$, we want a method for calculating the distribution of $g(X,Y)$ for some real valued function g of two variables, e.g. the sum: $g(x,y) = x+y$, product $g(x,y)=xy$, quotient, etc.

In general we will apply a density transformation method to a one-to-one transformation of the plane $u = g(x,y)$, $v = h(x,y)$ as covered by Proposition A, p. 102. For this method we find the joint density $f_{U,V}(u,v)$ of the pair $U := g(X,Y)$ and $V := h(X,Y)$ in terms of the transformation $(g; h)$ and the original joint density $f(x,y) = f_{X,Y}(x,y)$. One then finds the density of U by calculating the marginal density of the joint density $f_{UV}(u,v)$. Thus in the method of Proposition A one must introduce a second real valued transformation $v = h(x,y)$ that may or may not be of interest but at least provides a nice one-to-one transformation of the plane when paired with the original mapping g. We will illustrate below. But first, if g is particularly simple we may be able to apply the distribution function method directly <u>without</u> introducing a second real valued function $h(x,y)$.

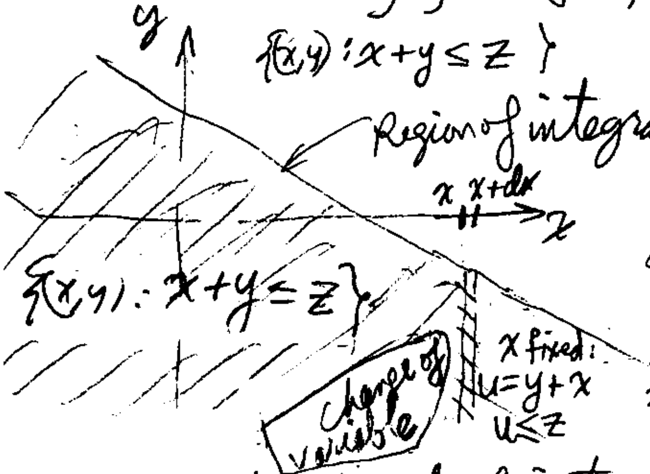
(over)

Notes 6.

p-2 <u>Distribution function method for calculating the density of $Z := X + Y$.</u>

Consider $Z := X + Y$. We have $F_Z(z) = P(X+Y \le z)$

$$= \iint_{\{(x,y):x+y\le z\}} f(x,y)\,dy\,dx, \quad \text{for each fixed } z \in (-\infty,\infty)$$



Region of integration for the double integral; <u>z is fixed.</u>

So, $F_Z(z) = \int_{-\infty}^{\infty}\int_{-\infty}^{z-x} f(x,y)\,dy\,dx$

To eliminate the dependence on x in the limit of integration, we make a change of variable for each x fixed on the variable y of the inner integral: $u = y + x$; then $du = dy$, and $y = u - x$ and finally u goes from $-\infty$ to z.

$$F_Z(z) = \int_{-\infty}^{\infty}\int_{-\infty}^{z} f(x, u-x)\,du\,dx$$

Now, since the region of integration in the double integral is a rectangle, we switch the order of integration to find

$$F_Z(z) = \int_{-\infty}^{z}\int_{-\infty}^{\infty} f(x, u-x)\,dx\,du.$$

Differentiating w.r.t. z finally, by the Fund. Thm of Calculus we have

$$(*) \qquad \boxed{f_Z(z) = \int_{-\infty}^{\infty} f(x, z-x)\,dx}.$$

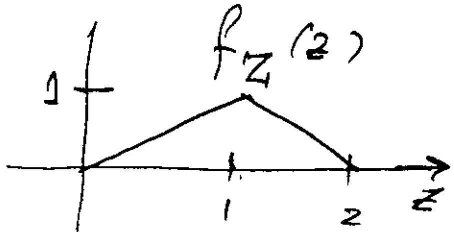
Example. Let $f(x,y)=\begin{cases}1 & 0\le x\le 1,\ 0\le y\le 1\\ 0 & \text{else}.\end{cases}$

Then $f_Z(z)=\int_{-\infty}^{\infty} f(x,z-x)\,dx = \int_{\{x:\,0\le x\le 1,\ 0\le z-x\le 1\}} 1\,dx$

$\Leftrightarrow \{x: 0\le x\le 1,\ x\le z,\ x\ge z-1\}$

$$=\begin{cases}\int_0^z 1\,dx, & 0\ z\le 1\\ \int_{z-1}^{1} 1\cdot dx, & 1\le z\le 2.\end{cases}$$

$$\therefore f_Z(z) = \begin{cases} z-0=z, & 0\le z\le 1.\\ 1-(z-1)=2-z, & 1\le z\le 2.\end{cases}$$



density of the sum of two independent Unif(0,1) r.v.'s

We apply Proposition A to recover the result (*), as follows. Define $U:=X+Y$ and $V:=X-Y$. This gives a one-to-one transformation of the plane.

We have that $f_{U,V}(u,v)\,dA = f_{X,Y}(x,y)\,dx\,dy$

where $dA = \left|\frac{\partial(u,v)}{\partial(x,y)}\right| dx\,dy$ for $\left|\frac{\partial(u,v)}{\partial(x,y)}\right| = \left|\det\begin{bmatrix}\frac{\partial u}{\partial x} & \frac{\partial u}{\partial y}\\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y}\end{bmatrix}\right|$

$= \left|\det\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}\right| = |-2| = 2.$ Therefore,

$$f_{U,V}(u,v) \underset{\text{Prop. A}}{=} f_{X,Y}(x,y)\frac{1}{\left|\frac{\partial(u,v)}{\partial(x,y)}\right|} = f_{X,Y}(x,y)\cdot\frac{1}{2}$$

where $x=\frac{u+v}{2}$, $y=\frac{u-v}{2}$. Therefore

$$f_{U,V}(u,v) = \frac{1}{2} f\left(\frac{u+v}{2}, \frac{u-v}{2}\right)$$

$\boxed{u:\text{fixed},\ x:=\frac{u+v}{2},\ dx=\frac{dv}{2} \text{ or } dv=2dx}$

Hence $f_U(u) = \int_{-\infty}^{\infty} \frac{1}{2} f\left(\frac{u+v}{2}, \frac{u-v}{2}\right) dv = \int_{-\infty}^{\infty} \cancel{\tfrac{1}{2}} f(x, u-x)\,\cancel{2}\,dx$

Notes 6

(p.4) 3.6.1. *Example A* Let $f(x,y) = \begin{cases} \lambda e^{-\lambda x} \cdot \lambda e^{-\lambda y}, & x,y>0, \\ 0, & \text{else.} \end{cases}$

be the joint density of two independent exponential r.v.'s each with parameter λ. Then by (*), or by the bottom of p. 97 (the convolution of the two exponential densities) we have, for $Z := X+Y$, that, for $z>0$,

$$f_Z(z) = \int_{\{x:\ 0<x<\infty,\ z-x>0\}} \lambda^2 e^{-\lambda x} e^{-\lambda(z-x)}\,dx$$

$$= \int_0^z \lambda^2 e^{-\lambda x} e^{-\lambda z+\lambda x}\,dx = \lambda^2 z e^{-\lambda z},\ z>0$$

This is the density of a Gamma r.v. with parameters λ and $\alpha=2$.

We can extend Example A, as follows:

Suppose $f_X(x) = \dfrac{\lambda^\alpha x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)}$, $x>0$ and $f_Y(y) = \dfrac{\lambda^\beta y^{\beta-1} e^{-\lambda y}}{\Gamma(\beta)}$, $y>0$

are Gamma densities with parameters λ and α and λ and β respectively. Let $Z := X+Y$ for independent r.v.'s X and Y with these respective Gamma densities, so again $f(x,y) = f_X(x) f_Y(y)$. The convolution of the two densities is a Gamma density with parameters λ and $\alpha+\beta$, as follows:

$$f_Z(z) = \int_0^z \frac{\lambda^{\alpha+\beta} x^{\alpha-1}(z-x)^{\beta-1} \cancel{e^{-\lambda x}} e^{-\lambda(z-\cancel{x})}}{\Gamma(\alpha)\Gamma(\beta)}\,dx$$

$$\theta = \frac{x}{z} \Rightarrow dx = z\,d\theta,\quad x^{\alpha-1} = z^{\alpha-1}\theta^{\alpha-1},\quad (z-x)^{\beta-1} = z^{\beta-1}(1-\theta)^{\beta-1}$$

$$\frac{\lambda^{\alpha+\beta} e^{-\lambda z}}{\Gamma(\alpha)\Gamma(\beta)} \int_0^z x^{\alpha-1}(z-x)^{\beta-1}\,dx = \frac{\lambda^{\alpha+\beta} e^{-\lambda z}}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 z^{\alpha+\beta-1}\theta^{\alpha-1}(1-\theta)^{\beta-1}\,d\theta$$

$$= \frac{\lambda^{\alpha+\beta} e^{-\lambda z} z^{\alpha+\beta-1}}{\Gamma(\alpha)\Gamma(\beta)} \cdot \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} = \boxed{\frac{\lambda^{\alpha+\beta} e^{-\lambda z} z^{\alpha+\beta-1}}{\Gamma(\alpha+\beta)},\ z>0}$$

Note 6.

(p.5) Note that the last computation automatically gives the result that $\int_0^1 \theta^{\alpha-1}(1-\theta)^{\beta-1}d\theta = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$

once we know the Gamma family. Indeed, the form $f_Z(z) = z^{\alpha+\beta-1} e^{-\lambda z} \lambda^{\alpha+\beta} \cdot C_{\alpha,\beta}, z \geq 0,$

for a constant $C_{\alpha,\beta}$, implies that $f_Z(z)$ is a Gamma density and that $C_{\alpha,\beta} = \frac{1}{\Gamma(\alpha+\beta)}$.

We apply this extension of 3.6.1 Example A as follows. Let X and Y be independent standard normal random variables ($X \sim N(0,1) \Rightarrow f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$)

Then first, the density of $U := X^2$ is obtained by proposition B, p. 62: $du = 2x\,dx$ or $\frac{dx}{du} = \frac{1}{2x}$

$$f_U(u) = \left(f_X(\sqrt{u}) + f_X(-\sqrt{u})\right)\left|\frac{dx}{du}\right|$$

$$= \left(\frac{1}{\sqrt{2\pi}} e^{-(\sqrt{u})^2/2} + \frac{1}{\sqrt{2\pi}} e^{-(-\sqrt{u})^2/2}\right) \frac{1}{2\sqrt{u}}$$

$$= \frac{1}{\sqrt{2\pi}} e^{-u/2} u^{-1/2}, u>0.$$ Since we know $\Gamma(\frac{1}{2}) = \sqrt{\pi}$,

this gives the Gamma density with $\lambda = 1/2$, $\alpha = 1/2$:

$$f_U(u) = \frac{(\frac{1}{2})^{1/2}}{\Gamma(1/2)} e^{-u/2} u^{1/2-1}, u>0.$$ Therefore,

if we put also $V := Y^2$ then $T = R^2 := X^2 + Y^2 = U + V$

has the Gamma density with parameters $\lambda = 1/2$ and $\alpha = \frac{1}{2} + \frac{1}{2} = 1$, that is an exponential density with $\lambda = \frac{1}{2}$: $f_T(t) = \frac{(\frac{1}{2})^1}{\Gamma(1)} e^{-t/2} t^0 = \frac{1}{2} e^{-t/2}$, $t > 0$.

~~Notes 6~~ (p. 6) 3.6.2 Example A.

If we introduce the other polar variable $\Theta = \arctan\left(\frac{Y}{X}\right)$ then we have $X = R\cos\Theta$, $Y = R\sin\Theta$ and we can find the joint density of R and Θ by Proposition A, p. 102:

$$f_{R,\Theta}(r,\theta) = f_{X,Y}(x,y)\left|\frac{\partial(x,y)}{\partial(r,\theta)}\right|$$

$$= \frac{e^{-(r\cos\theta)^2/2}}{\sqrt{2\pi}}\,\frac{e^{-(r\sin\theta)^2/2}}{\sqrt{2\pi}}\left|\det\begin{pmatrix}\cos\theta & -r\sin\theta\\ \sin\theta & r\cos\theta\end{pmatrix}\right|$$

$$= \frac{1}{2\pi}e^{-r^2/2}\cdot r. \qquad r>0,\quad 0<\theta<2\pi.$$

$$= \left\{\begin{matrix} re^{-r^2/2}, & 0<r<\infty \\ 0 & \text{else}\end{matrix}\right\}\cdot\left\{\begin{matrix}\frac{1}{2\pi}, & 0<\theta<2\pi \\ 0, & \text{else}\end{matrix}\right\}.$$

$$= f_R(r)\,f_\Theta(\theta) \quad \text{, a product of two densities}$$

This gives the result that Θ is Unif$(0,2\pi)$ and independent of R (!). Thus since also $T = R^2 \sim$ exponential$\left(\frac{1}{2}\right)$, one may generate a pair of independent standard normal variables as

$$X = \sqrt{T}\cos(\Theta),\quad Y = \sqrt{T}\sin(\Theta))$$

or $X = \sqrt{-2\ln(1-U_1)}\cos(2\pi U_2)$, $Y = \sqrt{-2\ln(1-U_1)}\sin(2\pi U_2)$ for independent Uniform$(0,1)$ r.v.s U_1 and U_2.