Math 621 Lec 8 9/26/17

Midterm 1

Midterm 2

Transformations of Variables

I Discrete r.v.'s

$X \sim \text{Bern}(p) = p^x(1-p)^{1-x} \mathbb{1}_{x\in\{0,1\}} = p_X(x)$

$Y = 3 + X \sim \begin{cases} 4 \text{ w.p. } p \\ 3 \text{ w.p. } 1-p \end{cases} = p^{y-3}(1-p)^{1-(y-3)} \mathbb{1}_{y\in\{3,4\}} = p_Y(y)$

$4-y$

$\text{Supp}(Y) = \{y : y-3 \in \text{Supp}(X)\}$

the PMF of Y looks like the PMF of X except x is replaced with y-3.

$Y = c + aX$

We have to keep the prob's of the original support values the same, but change the support.

$\Rightarrow$ if y was realized, that means $x = \frac{y-c}{a}$, the underlying initial realization

$\text{Supp}(Y) = \{y : \frac{y-c}{a} \in \text{Supp}(X)\} = \{y : \frac{y-c}{a} \in \{0,1\}\} = \{c, a+c\}$

$Y \sim p^{\frac{y-c}{a}}(1-p)^{1-\frac{y-c}{a}} \mathbb{1}_{y\in\{c,a+c\}}$

If $Y = c + aX = g(X)$ a transformation of X

$\Rightarrow g^{-1}(Y) = \frac{Y-c}{a}$

$\Rightarrow p_Y(y) = p^{g^{-1}(y)}(1-p)^{1-g^{-1}(y)} \mathbb{1}_{g^{-1}(y)\in\text{Supp}(X)} = p_X(g^{-1}(y))$

(working theory)

$X \sim \text{Bin}(n,p)$

$Y = a + cX \qquad p_Y(y) = \binom{n}{g^{-1}(y)} p^{g^{-1}(y)}(1-p)^{n-g^{-1}(y)} \mathbb{1}_{y\in g(\text{Supp}(X))}$

g's Image of A $\longrightarrow$ Note: $g(A) = \{g(a) : a \in A\}$

$$= \binom{n}{\frac{y-c}{a}} p^{\frac{y-c}{a}} (1-p)^{n-\frac{y-c}{a}} \mathbb{1}_{y \in \{c,\, a+c,\, 2a+c,\, \dots\, na+c\}} \qquad \frac{y-c}{a} = 0 = 1 = 2$$

$X \sim \text{Bin}(n,p)$, $Y = X^3$

$$p_Y(y) = \binom{n}{\sqrt[3]{y}} p^{\sqrt[3]{y}} (1-p)^{1-\sqrt[3]{y}} \mathbb{1}_{y \in \{0, 1, 2^3, 3^3, \dots, n^3\}}$$

$X \sim \text{Geom}(p)$

$Y = \max\{3, X\}$

What does this look like?

X	Y
0	3
1	3
2	3
3	3
4	4
5	5
⋮	⋮

There is no function $g^{-1}(y)$ because g is not 1:1.

First note $p_Y(4) = p_X(4)$, $p_Y(5) = p_X(5)$, ...

But $p_Y(3) \neq p_X(3)$

$$p_Y(3) = p_X(0) + p_X(1) + p_X(2) + p_X(3)$$

$\Rightarrow p_Y(y) \neq p_X(g^{-1}(y))$ always ... only for g functions which are 1:1 for support of X.

General formula:

$$p_Y(y) = \sum_{\{x : g(x) = y\}} p_X(x) \overset{\text{if } g \text{ 1:1}}{=} \sum_{\{x : x = g^{-1}(y)\}} p_X(x) = p_X(g^{-1}(y))$$

only one value!

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

Here, $p_Y(y) = \left(p + (1-p)p + (1-p)^2 p + (1-p)^3 p\right) \mathbb{1}_{y=3} + \underbrace{p(1-p)^y}_{\text{Geom}(p)} \mathbb{1}_{y \in \{4,5,\dots\}}$

$$F_Y(y) = \sum_{\{x : g(x) \le y\}} p_X(x)$$

$X_1, X_2 \overset{iid}{\sim} \text{Poisson}(\lambda)$, $Y = -X_2 \Rightarrow p_Y(y) = p_X(-y) = \dfrac{e^{-\lambda}\lambda^{-y}}{(-y)!} \mathbb{1}_{y \in \{0,-1,-2,\ldots\}}$

$D = X_1 - X_2 = X_1 + Y$ $\quad$ $\text{Supp}[D] = \mathbb{Z}$ (rare!)

$$p_D(d) = \sum_{x \in \text{Supp}[X_1]} p_{X_1}(x)\, p_Y(d-x) = \sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^x}{x!} \frac{e^{-\lambda}\lambda^{-(d-x)}}{\underbrace{(-(d-x))!}_{(x-d)!}} \mathbb{1}_{d-x \in \{0,-1,-2,\ldots\}}$$

$\overbrace{\lambda^x \lambda^{-d}}$

$d - x \in \{0,-1,-2,\ldots\}$
$\Updownarrow$
$x - d \in \{0,1,2,\ldots\}$
$\Updownarrow$
$x \in \{d, d+1, \ldots\}$

If $d > 0$, sum begins at d, if $d \le 0$, sum begins at 0 $\Rightarrow \max\{0,d\}$

$$= e^{-2\lambda} \begin{cases} \displaystyle\sum_{x=d}^{\infty} \frac{\lambda^{2x-d}}{x!(x-d)!} & \text{if } d > 0 \\ \displaystyle\sum_{x=0}^{\infty} \frac{\lambda^{2x-d}}{x!(x-d)!} & \text{if } d \le 0 \end{cases}$$

$\rightarrow$ let $x' = x - d \Rightarrow x = x' + d \Rightarrow \displaystyle\sum_{x'=0}^{\infty} \frac{\lambda^{\overbrace{2(x'+d)-d}^{\overbrace{2x'+2d-d}^{2x'-d}}}}{(x'+d)!\, x'!} = \underbrace{\sum_{i=0}^{\infty} \frac{\left(\frac{2\lambda}{2}\right)^{2i-d}}{\Gamma(i+d-1)\Gamma(i-1)}}$

modified Bessel Function of the 1st kind (sol'n to famous diff eq) denoted $I_d(2\lambda)$

$\rightarrow$ let $d' = -d \Rightarrow \displaystyle\sum_{x=0}^{\infty} \frac{\lambda^{2x+d'}}{x!(x+d')!} = \underbrace{\sum_{i=0}^{\infty} \frac{\left(\frac{2\lambda}{2}\right)^{2i-d'}}{\Gamma(i+d'-1)(i-1)}}_{I_{d'}(2\lambda)}$

if $d < 0 \Rightarrow d' = |d|$
if $d > 0 \Rightarrow d = |d|$

$\Rightarrow p_D(d) = e^{-2\lambda} I_{|d|}(2\lambda)$ = Skellam(λ, λ) distr. (1946)

Used to model point spreads in baseball, soccer, hockey, ..., difference in photon noise, and more.

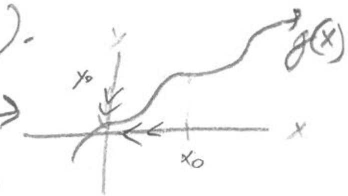
let $X \sim U(0,1)$, $Y = aX + c = g(X)$ s.t. g is 1:1

Can we use the formula

$P_Y(y) = P_X(g^{-1}(y))$? No! there is no $p_X(x)$ (PMF)! This does not generalize for cont. r.v.'s!!

We need another way!

Consider $Y = g(X)$ where g is 1:1. Find $f_Y(y)$ the PDF if $f_X(x)$ is known. If it's 1:1 it's either strictly increasing or strictly decreasing (thm from calculus).



(a) If g is increasing...

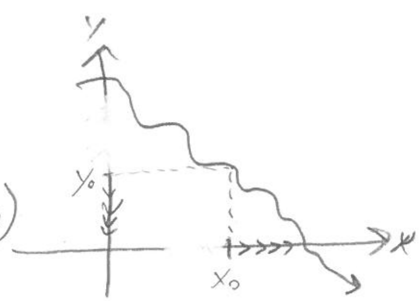
$$F_Y(y) := P(Y \le y) = P(g(X) \le \underbrace{y}_{y_0}) = P(X \le \underbrace{g^{-1}(y)}_{x_0}) = F_X(g^{-1}(y))$$

We just found the CDF of Y. Now for the PDF of Y:

$$f_Y(y) = F_Y'(y) = \frac{d}{dy}\left[F_X(g^{-1}(y))\right] = F_X'(g^{-1}(y))\frac{d}{dy}\left[g^{-1}(y)\right] = f_X(g^{-1}(y))\frac{d}{dy}\left[g^{-1}(y)\right]$$

by chain rule

(b) If g is decreasing,



$$F_Y(y) = P(Y \le y) = P(g(X) \le \underbrace{y}_{y_0}) = P(X \ge \underbrace{g^{-1}(y)}_{x_0})$$
$$= 1 - F_X(g^{-1}(y))$$

We have the CDF of Y. Now we need the PDF of Y:

$$\underbrace{f_Y(y)}_{\ge 0} = F_Y'(y) = \frac{d}{dy}\left[1 - F_X(g^{-1}(y))\right] = - \underbrace{f_X(g^{-1}(y))}_{\ge 0}\underbrace{\frac{d}{dy}\left[g^{-1}(y)\right]}_{\Rightarrow\ \le 0}$$

negative

Let's combine formulas (a) & (b) into one formula for convenience:

Note:

$$f_Y(y) = f_X\left(g^{-1}(y)\right)\left|\frac{d}{dy}\left[g^{-1}(y)\right]\right| \quad . \quad Supp[Y] = g(Supp(X)) = \{g(x) : x \in Supp(X)\} = \{y : g^{-1}(y) \in Supp(X)\}$$

If $Y = g(X) = aX + c$, $a, c \in \mathbb{R}$ s.t. $a \neq 0$, the linear transformation... then this becomes:

$$\Rightarrow y = ax + c \Rightarrow y - c = ax \Rightarrow x = \frac{y-c}{a} = g^{-1}(y) \qquad \left|\frac{d}{dy}\left[g^{-1}(y)\right]\right| = \frac{1}{|a|}$$

$$\Rightarrow f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-c}{a}\right)$$

Common linear transforms:

If $Y = -X \Rightarrow f_Y(y) = f_X(-y)$

If $Y = X + c \Rightarrow f_Y(y) = f_X(y - c)$

$X \sim U(0,1)$ & $Y = aX + c$

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-c}{a}\right) = \frac{1}{a}(1) = \frac{1}{|a|} \qquad Supp[Y] = [c, a+c] \Rightarrow Y \sim U(c, a+c)$$

$X \sim Exp(\lambda)$ & $Y = aX$ (No c right now) $\quad Supp(Y) = (0, \infty)$ if $a > 0$

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y}{a}\right) = \frac{1}{|a|}\lambda e^{-\lambda\frac{y}{a}} = \underbrace{\frac{\lambda}{|a|}}_{\lambda'} e^{-\overbrace{\frac{\lambda}{|a|}}^{\lambda'} y} = Exp\left(\frac{\lambda}{a}\right)$$

Can a be negative? No... since Supp of exp is $(0, \infty)$

It will be a r.v. just not an exponential

Special case:

$Y = X + c$

$\Rightarrow f_Y(y) = f_X(y - c)$ "shifted distribution"

$X \sim Exp(\lambda)$

$$f_Y(y) = \lambda e^{-\lambda(y-c)} = \left(e^{\lambda c}\right)\lambda e^{-\lambda y} \qquad Supp(Y) = (c, \infty)$$

scaling term in front $\qquad$ shift

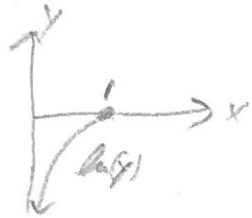
Prove $X \sim U(0,1) \Rightarrow Y = \underbrace{1}_{c}\underbrace{-}_{a=-1}X \sim U(0,1)$ makes sense!

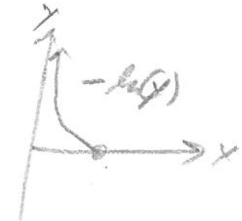
$$f_Y(y) = \frac{1}{|-1|} f_X\left(\frac{y-c}{-1}\right) = 1 = U(0,1)$$

$Supp(Y) = [0,1]$

$X \sim U(0,1)$, $Y = -\ln(X)$

$\Rightarrow \text{Supp}(Y) = (0,\infty)$





$f_Y(y) = f_X(g^{-1}(y))\left|\frac{d}{dy}(g^{-1}(y))\right|$ $\qquad y = -\ln(x) \Rightarrow x = e^{-y} = g^{-1}(y) \Rightarrow \left|\frac{d}{dy}(\;)\right| = e^{-y}$

$= (1)\,e^{-y} = \text{Exp}(1)$

$\Rightarrow Y = -\frac{1}{\lambda}\ln(X) \Rightarrow Y \sim \text{Exp}(\lambda)$ based on what we proved before

$\Rightarrow$ if you can generate $U(0,1)$'s, you can generate $\text{Exp}(\lambda)$'s easily!

$X \sim \text{Exp}(1)$ $\quad Y = -\ln\left(\frac{e^{-X}}{1-e^{-X}}\right)$ $\qquad$ What is Supp(Y)?

$= \ln\left(\frac{1-e^{-X}}{e^{-X}}\right)$

$= \ln(e^{X}-1)$

$X \in (0,\infty)$

$e^X \in (1,\infty)$

$\Rightarrow e^X - 1 \in (0,\infty)$

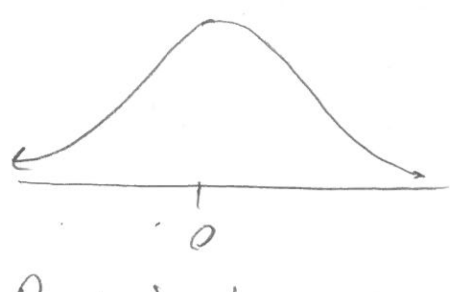
$\ln(e^X-1) \in (-\infty,\infty) \Rightarrow \text{Supp}(Y) = \mathbb{R}$

$f_Y(y) = f_X(g^{-1}(y))\left|\frac{d}{dy}(g^{-1}(y))\right|$

$y = \ln(e^x-1) \Rightarrow e^y = e^x - 1 \Rightarrow e^x = e^y + 1 \Rightarrow x = \underbrace{\ln(e^y+1)}_{g^{-1}(y)}$

$\left|\frac{d}{dy}[\;]\right| = \left|\frac{e^y}{e^y+1}\right| = \frac{e^y}{e^y+1}$ (always positive)

$f_Y(y) = e^{-\ln(e^y+1)}\frac{e^y}{e^y+1} = e^{\ln\left(\frac{1}{e^y+1}\right)}\frac{e^y}{e^y+1} = \frac{e^y}{(e^y+1)^2} = \text{Logistic}(0,1)$



Looks just like the normal but it has heavier tails

It is important in "deep learning" and "logistic regression".

Due to its heavier tails it is used for rating systems e.g. US Chess Federation. (Elo system)