Lec 9 MATH 340

Transformations of Discrete r.v.'s

$X \sim U(\{0,1,2,3\}) = \frac{1}{4}\mathbb{1}_{x \in \{0,1,2,3\}}$

let $Y = g(X) = -X$, a "transformation" of the r.v. $\iff X = -Y$

$Y \sim p_Y(y) = ?$

$p_Y(y) = P(Y=y) = P(-Y=-y) = P(X=-y) = p_X(-y) = \frac{1}{4}\mathbb{1}_{-y \in \{0,1,2,3\}}$

$= \frac{1}{4}\mathbb{1}_{y \in \{-3,-2,-1,0\}}$

$X \sim \text{Binom}(n,p) = \binom{n}{x} p^x (1-p)^{n-x}$

$Y = -X \sim \binom{n}{-y} p^{-y}(1-p)^{n+y}$ strange... but correct

---

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

$Y = \underbrace{X+3}_{g(X)} \sim p_Y(y) = ?$

$\Rightarrow X = Y-3$

$p_Y(y) = P(Y=y) = P(Y-3=y-3) = P(X=y-3) = p_X(y-3) = p^{y-3}(1-p)^{4-y}\mathbb{1}_{y \in \{3,4\}}$

What's the pattern? The use of the inverse function...

Assume g invertible...

$Y = g(X) \iff X = g^{-1}(Y)$

$p_Y(y) = P(Y=y) = P(g(X)=y) = P(X=g^{-1}(y)) = p_X(g^{-1}(y))$

$Y = aX+b$ shifted and scaled distr.

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

$\Rightarrow p_Y(y) = p_X(g^{-1}(y)) = p_X\left(\frac{y-b}{a}\right)$

$Y = -X \Rightarrow a=-1,\ b=0,\ p_Y(y) = p_X(-y)$

$Y = aX \Rightarrow b=0,\ p_Y(y) = p_X\left(\frac{y}{a}\right)$

$Y = X+b \Rightarrow a=1,\ p_Y(y) = p_X(y-b)$

$X_1, \ldots, X_n \overset{iid}{\sim} \text{Bern}(p)$

$\bar{X} = \frac{X_1 + \ldots + X_n}{n} = \frac{T}{n}$, $T \sim \text{Binom}(n, p)$

$\bar{X} \sim p_T(n\bar{x}) = \binom{n}{n\bar{x}} p^{n\bar{x}} (1-p)^{n(1-\bar{x})}$ distr. for Binomial Exact Test in 341

$X \sim \text{Binom}(n,p)$, $Y \overset{=g(X)}{=} X^2$ invertible? Generally not, but invertible on S_X and that's all that matters!

$\Downarrow$

$X = g^{-1}(Y) = \sqrt{Y}$

$Y \sim \binom{n}{\sqrt{y}} p^{\sqrt{y}} (1-p)^{n-\sqrt{y}}$

What if g is not invertible? Consider $X \sim U(\{1, 2, \ldots, 10\}) = \frac{1}{10} \mathbb{1}_{x \in \{1, 2, \ldots, 10\}}$

let $Y = g(X) = \min\{X, 3\}$. Let's figure this out manually...

Y	$p_Y(y)$
1	$\frac{1}{10}$
2	$\frac{1}{10}$
3	$\frac{8}{10}$

What did we do?

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

one value

$$p_Y(3) = \sum_{\{x : g(x) = 3\}} p_X(x) = \sum_{x \in \{3, 4, \ldots, 10\}} \frac{1}{10} = 8 \cdot \frac{1}{10}$$

For HW

e.g. $X \sim Bin(n,p)$, $Y = X^3$

$$p_Y(y) = \binom{n}{\sqrt[3]{y}} p^{\sqrt[3]{y}} (1-p)^{n-\sqrt[3]{y}}$$ looks weird... true!

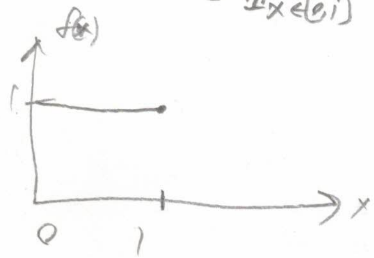
Transformations of Cont. r.v.'s

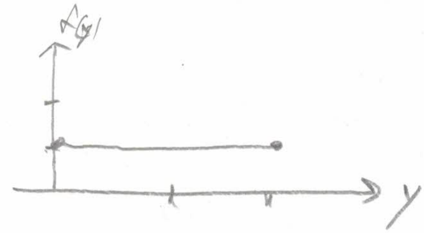
But does...

$$f_Y(y) \stackrel{?}{=} f_X(g^{-1}(y))$$

let $X \sim U(0,1)$, $Y = 2X = g(X)$, $\Rightarrow X = \frac{Y}{2} = g^{-1}(Y)$

$= \mathbb{1}_{x\in[0,1]}$

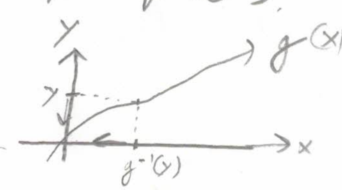




Likely $f_Y(y) = \frac{1}{2}\mathbb{1}_{y\in[0,2]}$

$f_Y(y) \stackrel{?}{=} f_X\left(\frac{y}{2}\right) = 1$ Something is wrong! Our discrete rule doesn't work for continuous r.v.'s since densities since densities are not probabilities. But CDF's represent prob's.

First consider g is a 1:1 function, strictly increasing (why??)



$$F_Y(y) = P(Y \le y) = P(g(X) \le y) = P(X \le g^{-1}(y)) = F_X(g^{-1}(y))$$

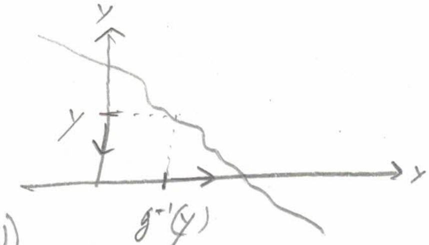
chain rule

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

$$= f_X(g^{-1}(y)) \frac{d}{dy}[g^{-1}(y)]$$ Note $\frac{d}{dy}[g^{-1}(y)] > 0$ by assumption

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

If g is 1:1 strictly decreasing



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

$f_Y(y) = -F_X'(g^{-1}(y)) \frac{d}{dy}[g^{-1}(y)]$ Note: $\frac{d}{dy}[g^{-1}(y)] < 0$ by assumption

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

Derive some rules! "Shifting and/or scaling"

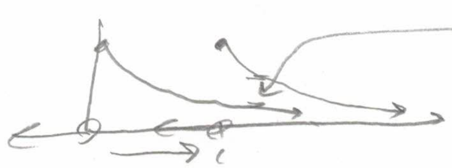
let $Y = g(X) = aX + b$ where $a, b \in \mathbb{R}$ constants but $a \neq 0$ o/w $Y \sim Deg(b)$

$\Rightarrow X = g^{-1}(Y) = \frac{Y-b}{a}$ $\left|\frac{d}{dy}[g^{-1}(y)]\right| = \frac{1}{|a|}$

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

If $Y = -X \Rightarrow a = -1, b = 0 \Rightarrow f_Y(y) = \frac{1}{|-1|} f_X(-y) = f_X(-y)$
If $Y = aX \Rightarrow b = 0 \Rightarrow f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y}{a}\right)$
If $Y = X + b \Rightarrow a = 1$, $f_Y(y) = f_X(y - b)$ shifted distr.

e.g. $X \sim Exp(\lambda) := \lambda e^{-\lambda x} \mathbb{1}_{x \in (0,\infty)} \Rightarrow f_Y(y) = \lambda e^{-\lambda(y-c)} \mathbb{1}_{y-c \in (0,\infty)}$
$= \underline{e^{\lambda c}} \lambda e^{-\lambda y} \mathbb{1}_{y \in (c,\infty)}$



constant multiple

$X \sim U(0,1)$, $Y = aX + b \sim ?$

if $a \ge 0$

$f_Y = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right) = \frac{1}{|a|} \mathbb{1}_{\frac{y-b}{a} \in [0,1]} = \frac{1}{|a|} \mathbb{1}_{y \in [b, b+a]}$

if $a < 0$ " $\frac{1}{|a|} \mathbb{1}_{y \in [b+a, b]}$

$Z \sim N(0,1)$, $X = \sigma Z + \mu \sim f_X(x) = ?$

Showing the cdf of X was hard... this way is easy...

$$f_X(x) = \frac{1}{|\sigma|} f_Z\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{|\sigma|}\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

We now have a new way to derive rvs: transform old ones!

$$X \sim \text{Exp}(1) = e^{-x}\mathbb{1}_{x\in(0,\infty)}$$

$$Y = -\ln\left(\frac{e^{-X}}{1-e^{-X}}\right) = \ln\left(\frac{1-e^{-X}}{e^{-X}}\right) = \ln(\underbrace{\underbrace{e^X}_{\in(1,\infty)}-1}_{\in(0,\infty)}) = g(X)$$, a 1:1 function

$$\Rightarrow e^Y = e^X - 1 \Rightarrow e^X = e^Y + 1 \Rightarrow X = \ln(e^Y+1) = g^{-1}(Y)$$

$$\left|\frac{d}{dx}\left[g^{-1}(Y)\right]\right| = \left|\frac{e^Y}{e^Y+1}\right| = \frac{e^Y}{e^Y+1}$$

$$f_Y(y) = f_X(g^{-1}(y))\left|\frac{d}{dy}\left[g^{-1}(y)\right]\right| = e^{-\ln(e^y+1)}\,\mathbb{1}_{\underbrace{\ln(e^y+1)\in(0,\infty)}}\,\frac{e^y}{e^y+1}$$

"std. logistic"

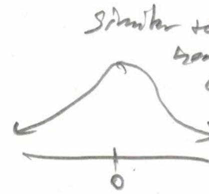
$e^y+1 \in (1,\infty)$

$e^y \in (0,\infty)$

$y \in (-\infty,\infty) = \mathbb{R}$

$$= \frac{1}{e^y+1}\,\frac{e^y}{e^y+1} = \frac{e^y}{(e^y+1)^2} = \text{Logistic}(0,1)$$

Similar to std normal curve



let $X = \sigma Y + \mu \sim \frac{1}{\sigma} f_Y\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{\sigma}\frac{e^{\frac{x-\mu}{\sigma}}}{\left(e^{\frac{x-\mu}{\sigma}}+1\right)^2} = \text{Logistic}(\mu,\sigma)$

where $\sigma > 0$, $\mu \in \mathbb{R}$, shift and scale transformation

used for chess ratings, logistic regression...

$$F_Y(y) = \int_{-\infty}^{y} \frac{e^t}{(e^t+1)^2}\,dt = \int_1^{1+e^y} \frac{e^{\ln(u-1)}}{\left(e^{\ln(u-1)}+1\right)^2}\,\frac{1}{u-1}\,du = \int_1^{1+e^y} \frac{u-1}{(u-1+1)^2}\,\frac{1}{u-1}\,du = \int_1^{1+e^y}\frac{1}{u^2}\,du = \left[-\frac{1}{u}\right]_1^{1+e^y} = 1 - \frac{1}{1+e^y}$$

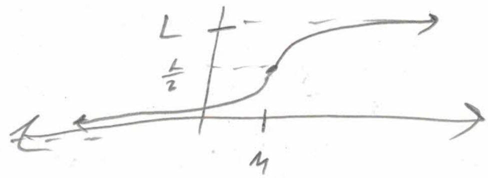
$$= \frac{e^y}{1+e^y} = \frac{1}{1+e}$$

let $u = 1 + e^t \Rightarrow u - 1 = e^t \Rightarrow t = \ln(u-1)$, $\frac{du}{dt} = e^t \Rightarrow dt = e^{-t}du = e^{-\ln(u-1)}du = e^{\ln\left(\frac{1}{u-1}\right)}du = \frac{1}{u-1}du$

if $t = -\infty \Rightarrow u = 1$, $t = y \Rightarrow u = 1 + e^y$

Define the "logistic function", a famous function

let $h(x) := \frac{L}{1+e^{-k(x-\mu)}}$ where
L: max value param
k: steepness param
μ: center param



if $L=1$, $k=1$, $\mu=0$

$h(x) = \frac{1}{1+e^{-x}}$, the "standard logistic function"

Since this is the CDF of r.v. Y, that's where it gets its name