MATH 340/640 LEC 18

Let X be a cont r.v. Let $Y = F_X(X)$. Weird!!

with support (a,b) where $a \in \mathbb{R}$ or $a = -\infty$, $b \in \mathbb{R}$ or $b = \infty$, $a < b$. All cont rv's we've seen satisfy this.

Y is a function of the rv X. The function is the CDF.

Is it 1:1? Yes! F is strictly increasing $\Rightarrow X = F_X^{-1}(Y)$

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

Inverse Functions:

let $u = g(t) \Longleftrightarrow t = g^{-1}(u)$

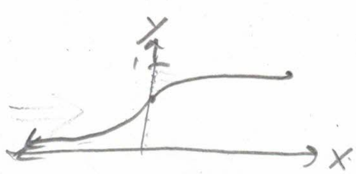
$$\frac{d}{du}[u] = \frac{d}{du}(g(t))$$

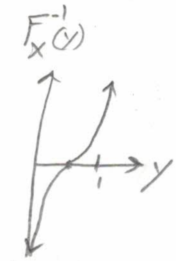
$$\Rightarrow 1 = g'(t)\frac{dt}{du} = g'(g^{-1}(u))\frac{d}{du}\left[g^{-1}(u)\right] \Rightarrow \frac{d}{du}\left[g^{-1}(u)\right] = \frac{1}{g'(g^{-1}(u))}$$

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

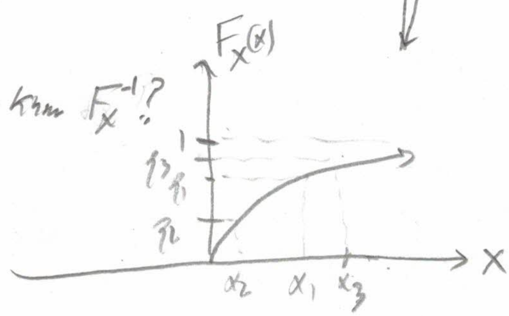
$$= \frac{f_X^{old}(F_X^{-1}(y))\mathbb{1}_{F_X^{-1}(y) \in S_X}}{f_X^{old}(F_X^{-1}(y))\mathbb{1}_{F_X^{-1}(y) \in S_X}} = \mathbb{1}^u_{F_X^{-1}(y) \in (a,b)} = \mathbb{1}^u_{y \in (F_X(a), F_X(b))} = \mathbb{1}^u_{y \in (0,1)} = U(0,1)$$

$\Rightarrow Y \sim U(0,1)$





(A1) How to sample x from $X \sim f(x)$, a cont. rv with known F_X^{-1}?



Step 1: Draw q from $U(0,1)$

Step 2: Compute $x = F_X^{-1}(q)$

For iid samples return to Step 1. Ensure the new q is independent from first q. Every computer language has a way to draw q from $U(0,1)$. This algorithm very fast! depending on how complicated F_X^{-1} is.

(A2) How to draw from $X \sim p(x)$, a discrete r.v. with known $F(x)$?

Step 1: Draw ξ from $U(0,1)$

Step 2: Compute $\min\{x : \text{s.t. } F(x) \geq \xi\}$

Can be made very fast via caching $F(x)$ values and using binary trees.

(A3) How to draw from X cont. if F_X^{-1} is unknown?..

(A3a) F_X is known, $S_X = (a,b)$ where a,b both finite. Only approx sampling is possible
in closed form

Select resolution Δ e.g. $\Delta = \frac{b-a}{1000}$

let $G = \{a, a+\Delta, a+2\Delta, \ldots, b\}$ ← grid

Compute $\{F(x) : x \in G\} = \{F(a), F(a+\Delta), F(a+2\Delta), \ldots, F(b)\}$. Use (A2). "Grid sampling"

(A3b) only $f(x)$ is known, same support. Select Δ, create G

The below should be summing up the Δf(x) terms just like a Riemann Integral rectangle approx.

Compute $\langle F(x) = \Delta f(x) : x \in G \rangle = \langle F(a) = \Delta f(a), F(a+\Delta) = \Delta f(a+\Delta), \ldots \Delta f(b) \rangle$ Riemann Integral. Use (A2)

(A3c) only $k(x)$ is known, same support. Select Δ

Compute $\langle k(x) : x \in G \rangle$

Compute $c = \frac{1}{\sum_{x \in G} k(x)}$. Compute $\langle f(x) = c\,k(x) : x \in G \rangle$. Use (A3b)

If $a = -\infty$, then you will have to use your best judgment to set a "negative enough" to not interfere with support. This involves having some prior knowledge about the r.v. you're sampling from.

If $b = \infty$, ,, ,, ,, "large enough" . . . . . .

(A4a) If X is discrete and $F(x)$ isn't known but $p(x)$ is ... HW

(A4b) ,, ,, ,, ,, ,, ,, $p(x)$ . . . . $k(x)$ . . . . . HW

Conditional Distributions and Mixture/Compound Distributions

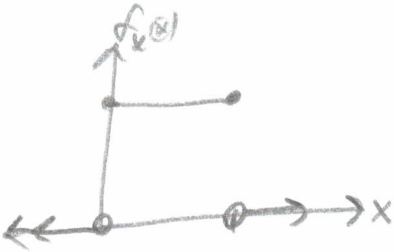
let $X \sim U(0,1)$, $Y|X=x \sim U(0,x)$, a "hierarchical" rv model

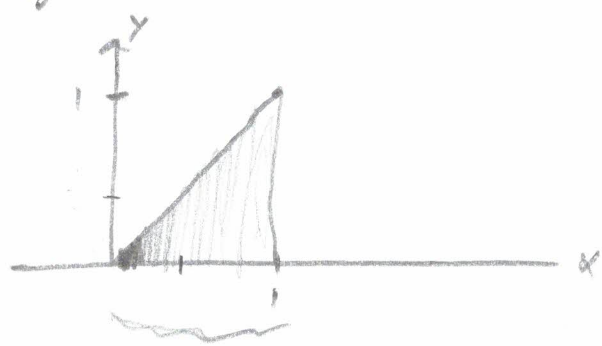
How is Y realized? X is realized first, then Y is realized from the value of x. Typical illustration

X | Y|X

e.g. $x = 0.3$
$Y \sim U(0, 0.3)$, $y = 0.1$

So how does the joint density look?





density is spread evenly over x but not over y as it has to "stretch" upwards

Can we find $f_{X,Y}(x,y)$? $f_Y(y)$? $f_{X|Y}(x,y)$? Yes... Recall

Marginalizing: $f_Y(y) = \int_R f_{X,Y}(x,y)\,dx$, $f_X(x) = \int_R f_{X,Y}(x,y)\,dy$ or with PMF's... replace $\int_R$ by $\sum_{x \in R}$ Bayes Rules:

Cond Probabilities: $f_{X|Y}(x,y) = \dfrac{f_{X,Y}(x,y)}{f_Y(y)} \Rightarrow f_{X,Y}(x,y) = f_{X|Y}(x,y)\,f_Y(y)$ $\Rightarrow$ $f_{X|Y}(x,y) = \dfrac{f_{Y|X}(x,y)\,f_X(x)}{f_Y(y)}$

$f_{Y|X}(x,y) = \dfrac{f_{X,Y}(x,y)}{f_X(x)} \Rightarrow f_{X,Y}(x,y) = f_{Y|X}(x,y)\,f_X(x)$ $\Rightarrow$ $f_{Y|X}(x,y) = \dfrac{f_{X|Y}(x,y)\,f_Y(y)}{f_X(x)}$

Bayes Thms: Marginal + Bayes Rules:

$$\Rightarrow \quad f_{X|Y}(x,y) = \frac{f_{Y|X}(x,y)\, f_X(x)}{\int_{\mathbb{R}} f_{Y|X}(x,y)\, f_X(x)\, dx}, \quad f_{Y|X}(x,y) = \frac{f_{X|Y}(x,y)\, f_Y(y)}{\int_{\mathbb{R}} f_{X|Y}(x,y)\, f_Y(y)\, dy}$$

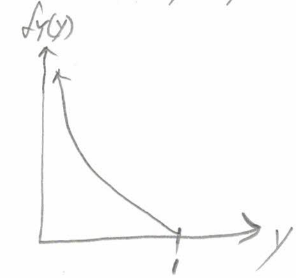
Let's now derive $f_{X,Y}(x,y)$, $f_Y(y)$, $f_{X|Y}(x,y)$. We know

$$f_X(x) = \mathbb{1}_{x\in(0,1)}, \quad f_{Y|X}(x,y) = \frac{1}{x}\mathbb{1}_{y\in[0,x]}$$

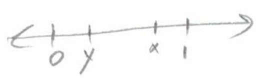
$$\Rightarrow f_{X,Y}(x,y) = \frac{1}{x}\mathbb{1}_{y\in[0,x]}\mathbb{1}_{x\in(0,1)}$$

$$f_Y(y) = \int_{\mathbb{R}} \frac{1}{x}\mathbb{1}_{y\in[0,x]}\mathbb{1}_{x\in[0,1]}\,dx = \int_0^1 \frac{1}{x}\mathbb{1}_{y\in(0,1)}\mathbb{1}_{x\in[y,1)}\,dx = \mathbb{1}_{y\in(0,1)}\int_y^1 \frac{1}{x}\,dx = \ln(x)\Big]_y^1 \mathbb{1}_{y\in(0,1)}$$

$$= -\ln(y)\,\mathbb{1}_{y\in(0,1)}$$



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

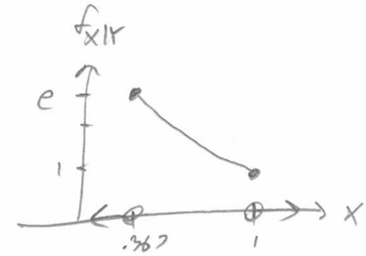


$$f_{X|Y}(x,y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{\frac{1}{x}\mathbb{1}_{y\in(0,1)}\mathbb{1}_{x\in[y,1)}}{-\ln(y)\,\mathbb{1}_{y\in(0,1)}} = -\frac{1}{x\ln(y)}\mathbb{1}_{x\in[y,1)}\,\mathbb{1}''_{y\in(0,1)}$$

What does this look like?

$$f_{X|Y}(x, e^{-1}) = -\frac{1}{x\ln(e^{-1})}\mathbb{1}_{x\in[e^{-1},1]} = \frac{1}{x}\mathbb{1}_{x\in[e^{-1},1)}$$

$\frac{1}{e} \approx .367$

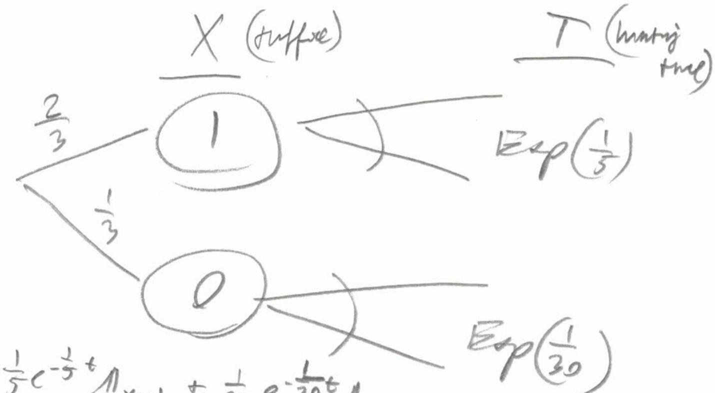


Here, $Y \sim f_Y(y)$ is known as a compound distribution as the "components" of Y, the $U(0,x)$'s were compounded together by the continuous distribution $X \sim f_X(x)$.

Let's do an example we can more easily wrap our head around:

Mixture Distribution

Consider a generalization of the exponential waiting time. Here, $\frac{2}{3}$ of the time there is no traffic you wait $\text{Exp}\left(\frac{1}{5}\right)$ sec and $\frac{1}{3}$ of the time there is traffic you wait $\text{Exp}\left(\frac{1}{30}\right)$ sec, let's draw an illustration:



Let's "create" the mixture variable $X = \begin{cases} 1 & \text{if no traffic} \\ 0 & \text{if traffic} \end{cases}$ which is $\text{Bern}\left(\frac{2}{3}\right)$ distributed

$\Rightarrow T \mid X=1 \sim \text{Exp}\left(\frac{1}{5}\right)$

$T \mid X=0 \sim \text{Exp}\left(\frac{1}{30}\right)$

$$\Rightarrow f_{T|X}(t,x) = \frac{1}{5} e^{-\frac{1}{5}t} \mathbb{1}_{x=1} + \frac{1}{30} e^{-\frac{1}{30}t} \mathbb{1}_{x=0}$$

What is the "marginal" waiting time i.e. $T \sim f_T(t) = ?$ Let's use our rules of

$$f_T(t) = \sum_{x \in \mathbb{R}} f_{T,X}(t,x) = \sum_{x \in \mathbb{R}} f_{T|X}(t,x)\, p_X(x) = \sum_{x \in \{0,1\}} f_{T|X}(t,x) \left(\frac{2}{3}\right)^x \left(\frac{1}{3}\right)^{1-x} \mathbb{1}_{x \in \{0,1\}}$$

$$= f_{T|X}(t,0)\, \frac{1}{3} + f_{T|X}(t,1)\, \frac{2}{3}$$

$$= \frac{1}{3} \frac{1}{30} e^{-\frac{1}{30}t} + \frac{2}{3} \cdot \frac{1}{5} e^{-\frac{1}{5}t}$$

If you waited 25 sec, what is the prob of traffic?

Bayes Rule...

$$p_{X|T}(x,t) = \frac{f_{T|X}(t,x)\, p_X(x)}{f_T(t)} = \frac{\left(\frac{1}{5}e^{-\frac{1}{5}t}\mathbb{1}_{x=1} + \frac{1}{30}e^{-\frac{1}{30}t}\mathbb{1}_{x=0}\right)\left(\frac{2}{3}\right)^x\left(\frac{1}{3}\right)^{1-x}\mathbb{1}_{x\in\{0,1\}}}{\frac{1}{3}\frac{1}{30}e^{-\frac{1}{30}t} + \frac{2}{3}\frac{1}{5}e^{-\frac{1}{5}t}}$$

$$p_{X|T}(0,25) = \frac{\frac{1}{3}\frac{1}{30}e^{-\frac{1}{30}(25)}}{\frac{1}{3}\frac{1}{30}e^{-\frac{1}{30}(25)} + \frac{2}{3}\frac{1}{5}e^{-\frac{1}{5}(25)}} = \frac{0.00483}{0.00483 + 0.000898} = .843$$

Reasonable $E[T \mid X=0] = 30$, $E[T \mid X=1] = 5$. 25 much closer to .

Given of waiting time longer, likely traffic.

The denominator in Bayes thm is called a mixture or compound distr:

$$p(y) = \sum_{x \in R} p_{Y|X}(x,y)\, p(x)$$ if Y is discrete, X is discrete

$$p(y) = \int_R p_{Y|X}(x,y)\, f(x)\, dx$$ if Y is discrete X is continuous

$$f(y) = \sum_{x \in R} f_{Y|X}(x,y)\, p(x)$$ if Y is cont. and X is discrete

$$f(y) = \int_R f_{Y|X}(x,y)\, f(x)\, dx$$ if Y is cont and X is cont.

"mixture distr" if X is discrete or "compound distr" if X is cont.

"component" distr.

"weight" distr or "mixing" distr