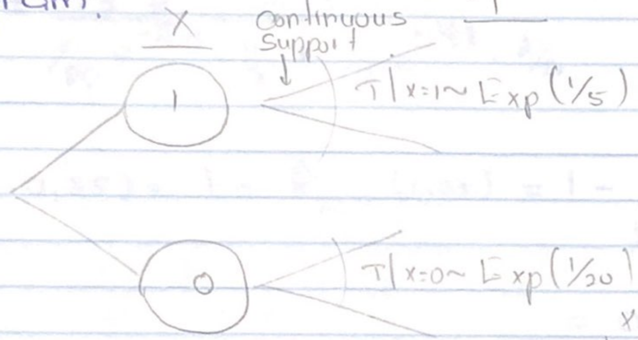
Lecture - 15 10/28/2020

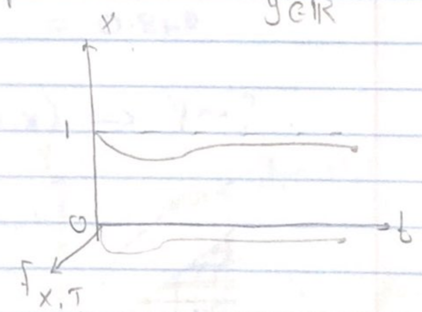
Mixture and compound distribution

Consider a situation where 2/3 of the time there is fast internet speed so your downloads Take $T \sim Exp(1/5) \Rightarrow E[T] = 5s$ and the other 1/3 of the time, there is Internet traffic, so your downloads take $T \sim Exp(1/20) \Rightarrow E[T] = 20s$. What is the distribution of the "overall T" or "unconditional on the Internet speed? Let $X \sim Bern(2/3)$ and $X = 1$ corresponds to fast internet and $X = 0$ corresponds to slow internet. Let's draw a tree diagram:



$$h(x) = \int_{\mathbb{R}} h(x,y)\,dy$$

$$h(x) = \sum_{y \in \mathbb{R}} h(x,y)$$



$$f_T(t) = \sum_{x \in Supp[X]} f_{T,X}(t,x)$$

$$= \sum_{x \in Supp[X]} f_{T|X}(t,x)\, P_X(x)$$

$$= \sum_{x \in \{0,1\}} f_{T|X}(t,x)\, p_X(x) = f_{T|X}(t,0)\, P_X(0) + f_{T|X}(t,1)\, P_X(1)$$

$$= \frac{1}{20} e^{-\frac{1}{20}t} \cdot 1/3 + \frac{1}{5} e^{-\frac{1}{5}t} \cdot 2/3$$

If the download speed was $t = 25s$, what is the probability it is a slow internet day, i.e. $X = 0$?

$$X | T \sim \text{Bern}(?)$$

$$P_{X|T}(x,t) = \frac{f_{T|X}(t,x)\, P_X(x)}{f_T(t)}$$ "Bayes Rule"

$W \sim \text{Bern}(p)$
$p = P(W=1)$

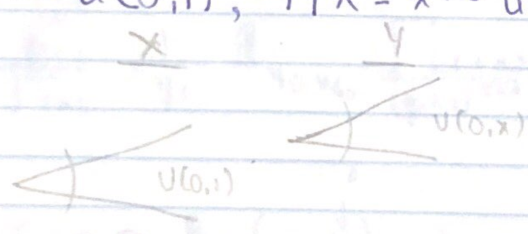
$$\text{Bernoulli param} = P_{X|T}(1,t) = \frac{f_{T|X}(t,1)\, P_X(1)}{f_T(t)}$$

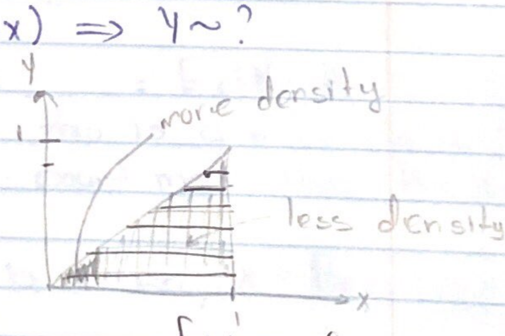
$$= \frac{\frac{1}{5} e^{-\frac{1}{5}t} \cdot \frac{2}{3}}{\frac{1}{20} e^{-\frac{1}{20}t} \cdot \frac{1}{3} + \frac{1}{5} e^{-\frac{1}{5}t} \cdot \frac{2}{3}}$$

$$P_{X|T}(0,25) = 1 - P_{X|T}(1,25) = 1 - \frac{\frac{1}{5} e^{-\frac{1}{5}\cdot 25} \cdot \frac{2}{3}}{\frac{1}{20} e^{-\frac{1}{20}\cdot 25} \cdot \frac{1}{3} + \frac{1}{5} e^{-\frac{1}{5}\cdot 25} \cdot \frac{2}{3}}$$

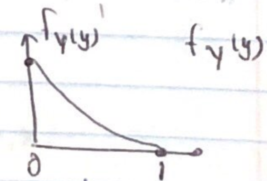
$$= 0.842$$

$$X \sim U(0,1),\ Y | X = x \sim U(0,x) \Rightarrow Y \sim ?$$





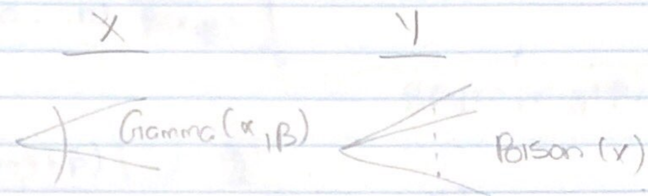
First example featured T which was continuous (we call that the 'model") and X which is discrete (we call that the "mixing distribution". ~~In the 2nd example Y,~~ Thus the unconditional distribution T is called a "Mixture distribution".



In the second example Y, the model is continuous and X, the mixing distribution is also continuous and we call the unconditional distribution Y a "compound distribution".

p 156–157 $\quad$ Let $Y|X=x \sim \text{Poisson}(x)$, $X \sim \text{Gamma}(\alpha, \beta)$

$Y \sim$?



$$P_Y(y) = \int_{\text{supp}[X]} P_{Y|X}(y,x)\, f_X^{\text{old}}(x)\, dx = \int_0^{\infty} \frac{e^{-x} x^y}{y!} \mathbb{1}_{y \in \mathbb{N}_0} \cdot \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}\, dx$$

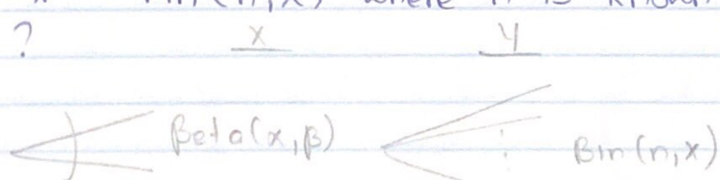
$$= \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{1}{y!} \mathbb{1}_{y \in \mathbb{N}_0} \underbrace{\int_0^{\infty} x^{y+\alpha-1} e^{-(\beta+1)x}\, dx}_{\text{lec 9 integral}}$$

$$= \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{1}{y!} \mathbb{1}_{y \in \mathbb{N}_0} \frac{\Gamma(x+\alpha)}{(\beta+1)^{x+\alpha}} = \_\_\_ = \text{ExtNegBin}\left(\alpha, \frac{\beta}{\beta+1}\right)$$

this is a more flexible count model than the Poisson.

$Y|X=x \sim \text{Bin}(n, \overset{p}{\downarrow}{x})$ where n is known, $X \sim \text{Beta}(\alpha, \beta)$

$Y \sim$?



$$P_Y(y) = \int_{supp[X]} P_{Y|X}(y,x)\, f_X^{old}(x)\, dx = \int_0^1 \binom{n}{y} x^y (1-x)^{n-y} \mathbb{1}_{y\in\{0,\dots,n\}} \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1} dx.$$

$$= \binom{n}{y} \mathbb{1}_{y\in\{0,\dots,n\}} \frac{1}{B(\alpha,\beta)} \underbrace{\int_0^1 x^{y+\alpha-1}(1-x)^{n-y+\beta-1} dx}_{B(y+\alpha,\, n-y+\beta)}$$

$$= \frac{B(y+\alpha, n-y+\beta)\binom{n}{y}}{B(\alpha,\beta)} \mathbb{1}_{y\in\{0,\dots,n\}}.$$

$$\text{II} \qquad \text{BetaBinomial}(n,\alpha,\beta)$$

$Y|X=x \sim \text{Exp}(x)$, $X \sim \text{Gamma}(\alpha,\beta) \overset{HW}{\Rightarrow} Y \sim \text{Lomax}(\beta,\alpha)$

Which is a more flexible waiting time than the exponential

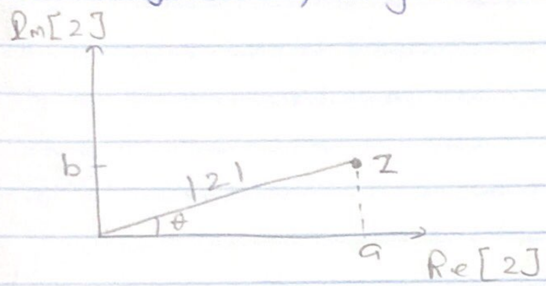
Midterm II ↑.

Final
⇓
Moment generating functions (mgf's) and characteristic functions (chf's). To derive these, we need to review complex/imaginary numbers. First define $i := \sqrt{-1}$ "imaginary"

let $a, b \in \mathbb{R}$, $z := a + bi \in \mathbb{C}$, complex #'s

$\mathrm{Re}[z] := a$, $\mathrm{Im}[z] := b$, real component and imaginary component of a complex #

$|z| := \sqrt{a^2 + b^2}$, $\mathrm{Arg}[z] := \theta \overset{\text{usually}}{=} \arctan(b/a)$



$i^0 = 1$
$i^1 = i$
$i^2 = (\sqrt{-1})^2 = -1$
$i^3 = i^2 i = -i$ clock cycle
$i^4 = (i^2)^2 = 1$
$i^5 = i^4 i = i \ldots n \in \mathbb{N}, i^n = i^{n \bmod 4}$

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \ldots$$

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \ldots$$

$$\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \ldots$$

$$e^{itx} = 1 + itx - \frac{t^2x^2}{2!} - \frac{it^3x^3}{3!} + \frac{t^4x^4}{4!} + \frac{it^5x^5}{5!} + \ldots$$

$$i\sin(tx) = itx - \frac{it^3x^3}{3!} + \frac{it^5x^5}{5!} + \ldots$$

$$\cos(tx) = 1 - \frac{t^2x^2}{2!} + \frac{t^4x^4}{4!} - \ldots$$

$$\Rightarrow e^{itx} = i\sin(tx) + \cos(tx) \xRightarrow{tx=\pi} e^{i\pi} = -1$$

$$\Rightarrow e^{i\pi} + 1 = 0.$$

Euler's Formula