BSCMA1004

STATISTICS II NOTES

WEEK 5 NOTES

IITM B.S Degree

PREPARED BY

Vehaan Handa, IIT Madras

Functions of Continuous RV's

Eg. Suppose $X \sim Unif[0,1]$

· $Y = 2X \in [0,2]$ What is distribution of Y?

So you have to first see if Y is $Unif[0,2]$.

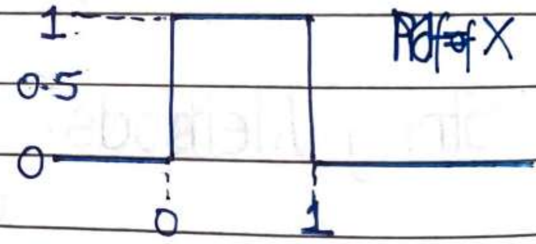
Check the CDF first :→

For $y \in [0,2]$

$$F_Y(y) = P(Y \leq y) = P(2X \leq y)$$
$$= P(X \leq y/2)$$
$$= \int_0^{y/2} f_X(x)\,dx = \frac{y}{2}$$

PDF of Y, $f_Y(y) = \frac{dF_Y(Y)}{dy} = \frac{1}{2}$

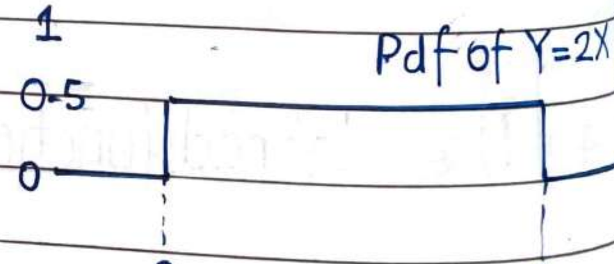
$Y \sim Uniform[0,2]$



In general if :→

$Y = aX + b$, then

$Y \sim Unif[b, b+a]$



DATE

CDF of $g(X)$:

- Suppose X is a CRV with cdf $F(x)$ and pdf f_X.
- Suppose $g : \mathbb{R} \to \mathbb{R}$ is a reasonable function.
- Then $Y = g(X)$ is a rv with CDF F_Y determined as:
  - $F_Y(y) = P(Y \leq y) = P(g(X) \leq y) = P(X \in \{x : g(x) \leq y\})$

Convert subset $A_Y = \{x : g(x) \leq y\}$ into intervals on real line.
Find probability that X falls in those intervals.
$F_Y(y) = P(X \in A_Y) = \int_{A_Y} f_X(x)\,dx$

If F_Y has no jumps, differentiating could help finding a PDF.

For monotonic, differentiable functions

Suppose X is a CRV with PDF f_X. Let $g(x)$ be monotonic for $x \in \text{supp}(X)$ with derivative $g'(x) = \frac{d}{dx} g(x)$

Then PDF of $Y = g(X)$ is :→

$$\boxed{f_Y(y) = \frac{1}{|g'(g^{-1}(y))|} f_X(g^{-1}(y))}$$

Translation: $Y = X + a$

$g(x) = x + a$, $g'(x) = 1$ $\quad y = x + a \Rightarrow x = y - a$

$\Rightarrow g^{-1}(y) = y - a$

$$\boxed{f_Y(y) = f_X(y - a)}$$

<u>Scaling</u>: $Y = aX$ $\Rightarrow x = y/a$

$(g(x) = ax,\ g'(x) = a,\ y = ax \Rightarrow g^{-1}(y) = y/a)$

$$\boxed{f_Y(y) = \frac{1}{|a|} f_X(y/a)}$$

<u>Affine</u>: $Y = aX + b$

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

★ <u>Affine transformation of Normal Distributions</u>

- $X \sim \text{Normal}(0,1)$ i.e standard normal.

$$f_X(x) = \frac{1}{\sqrt{2\pi}} \exp(-x^2/2)$$

- $Y = \sigma X + \mu$

$$f_Y(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-(x-\mu)^2/2\sigma^2\right)$$

$Y \sim \text{Normal}(\mu, \sigma^2)$

- $X \sim \text{Normal}(\mu, \sigma^2)$

$Y = \frac{(X-\mu)}{\sigma} \sim \text{Normal}(0,1)$

So affine transformation of a normal rv is normal.

DATE

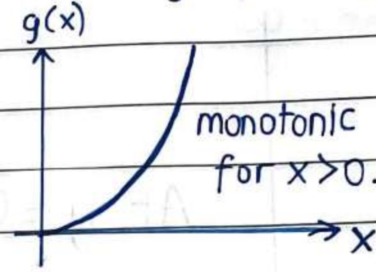
Q. Let $X \sim Exp(\lambda)$. Find pdf of X^2.

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & \text{otherwise} \end{cases} \qquad supp(X) = \{x : x > 0\}$$

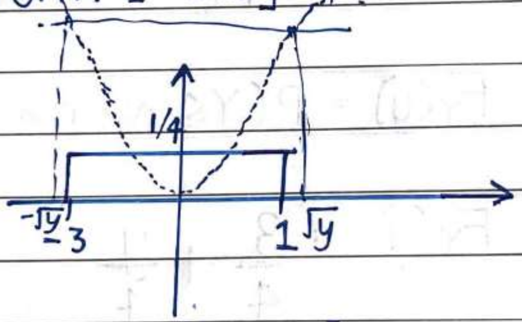
$Y = X^2 \quad g(x) = x^2, \; g'(x) = 2x \qquad y = x^2 \quad x = \sqrt{y}$

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

$$f_Y(y) = \frac{1}{2\sqrt{y}} \lambda e^{-\lambda\sqrt{y}} \qquad y > 0$$



Q. Let $X \sim$ Uniform $[-3,1]$. Find PDF of X^2.



$supp(X) = [-3,1]$

$g(x) = x^2$ is not monotonic $\because$ in $supp(X)$

$Y = X^2 \in [0,9] \qquad y \in [0,9]$

$F_Y(y) = P(Y \leq y) = P(X^2 \leq y)$ ~~$y \in$~~

Say $y \in [0,1]$: $(X^2 \leq y) \leftrightarrow -\sqrt{y} < x < \sqrt{y} \Rightarrow F_Y(y) = \frac{2\sqrt{y}}{4}$

$y \in [1,9]$: $(X^2 \leq y) \leftrightarrow -\sqrt{y} < x < 1 \Rightarrow F_Y(y) = \frac{1+\sqrt{y}}{4}$

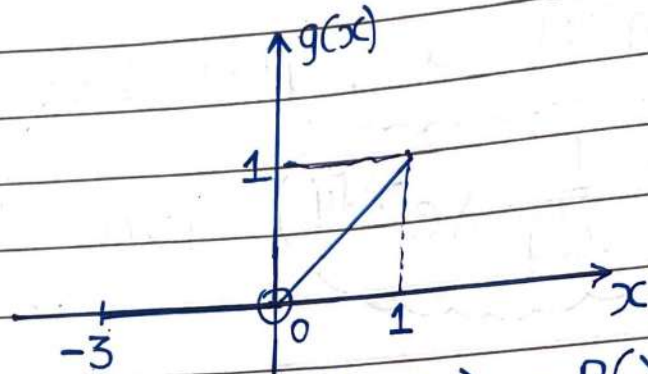
$$f_Y(y) = \begin{cases} \frac{1}{2} \cdot \frac{1}{2\sqrt{y}}, & 0 < y < 1 \\ \frac{1}{4} \cdot \frac{1}{2\sqrt{y}}, & 1 < y < 9 \end{cases}$$

DATE

Q. Let $X \sim \text{Uniform}[-3,1]$. Find PDF of $\max(X,0)$

$$g(x) = \max(x,0) = \begin{cases} 0, & \text{if } -3<x<0 \\ x, & \text{if } 0<x<1 \end{cases}$$

$Y = g(X) \in [0,1]$



At $y=0$ $\quad F_Y(y) = P(Y \leq 0) = P(Y=0)$ → Y is not continuous

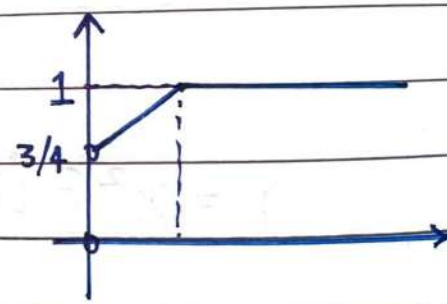
$= P(g(X) \leq 0)$

$= P(-3<X<0) = 3/4$

If $y<0$: $F_Y(y) = P(Y \leq \text{-ve num}) = 0$

$0<y<1$: $F_Y(y) = \dfrac{3}{4} + \dfrac{y}{4}$

$y>1$: $F_Y(y) = 1$



DATE

★ Expected value : Continuous RV's :

Let X be a CRV with density $f_X(x)$. Let $g: \mathbb{R} \to \mathbb{R}$ be a function. The expected value of $g(X)$, $E[g(X)]$ is given by :

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)\,dx$$

whenever the integral exists. It may diverge to $\pm\infty$ sometimes

★ Mean and Variance

X : continuous RV

- Mean, or $E[X]$ or μ_X or μ, when $g(x) = X$

$$E[X] = \int_{-\infty}^{\infty} x f(x)\,dx$$

- Variance, or $Var[X]$ or σ_X^2 or σ^2

$$Var(X) = E[(X-\mu_X)^2] = \int_{-\infty}^{\infty} (x-\mu_X)^2 f_X(x)\,dx$$

- Measure of spread abt mean
- $Var(X) = E[X^2] - \mu^2$

- $X \sim \text{Uniform}[a,b]$: $f_X(x) = \frac{1}{b-a}$, $a < x < b$
  - $E[X] = \frac{a+b}{2}$, $Var(X) = \frac{(b-a)^2}{12}$
- $X \sim \text{Exp}(\lambda)$: $f_X(x) = \lambda \exp(-\lambda x)$, $x > 0$, $E[X] = 1/\lambda$ $Var(X) = 1/\lambda^2$
- $X \sim \text{Normal}(\mu, \sigma^2)$, $f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$ $E[X] = \mu$, $Var(X) = \sigma^2$

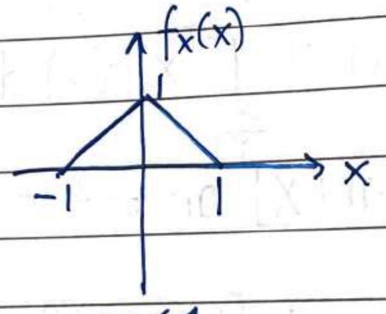
★ Probability Space: Continuous case

- Discrete case: → Sample space is finite with PMF
- Here S is an interval of real line
- Events: intervals in S along with complements and countable unions.
- Probability function: Fn from intervals inside S to [0,1] satisfying axioms. Possible for $P(X=x)=0$.

Q. CRV X has PDF: →

$$f_X(x) = \begin{cases} 1-|x|, & -1 \leq x \leq 1 \\ 0, & \text{otherwise} \end{cases}$$

Find CDF, $E[X]$, $Var(X)$.



$$F_X(x) = \begin{cases} 0, & x < 1 \\ & -1 \leq x < 0 \\ & 0 \leq x < 1 \\ 1, & x \geq 1 \end{cases}$$

$-1 \leq x < 0$ $\quad F_X(x) = \int_{-1}^{x} (1-|u|)\,du = \int_{-1}^{x} (1+u)\,du.$

$$= x+1-\frac{x^2}{2}-\frac{1}{2}$$

$$= \frac{1}{2}+x+\frac{x^2}{2}$$

discrete

DATE

$0 \leq x < 1$ $\quad F_X(x) = \underbrace{\int_{-1}^{0}(1-|u|)\,du}_{1/2} + \int_{0}^{x}(1-|u|)\,du$

$$= \frac{1}{2} + u\Big|_0^x - \frac{u^2}{2}\Big|_0^x = \frac{1}{2} + x - \frac{x^2}{2}$$

$$E[X] = \int_{-1}^{1} x f_X(x)\,dx = \int_{-1}^{0} x(1+x)\,dx + \int_{0}^{1} x(1-x)\,dx$$

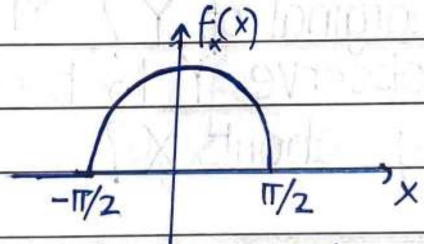
$$= \frac{x^2}{2}\Big|_{-1}^{0} + \frac{x^3}{3}\Big|_{-1}^{0} + \frac{x^2}{2}\Big|_{0}^{1} - \frac{x^3}{3}\Big|_{0}^{1} = \frac{-1}{2} + \frac{1}{3} + \frac{1}{2} - \frac{1}{3} = 0$$

$$Var(X) = E[X^2] = \frac{1}{6}$$

Q. CRV X has PDF:→

$$f_X(x) = \begin{cases} \frac{1}{2}\cos x \ ; & -\pi/2 \leq x \leq \pi/2 \\ 0 \ ; & \text{otherwise} \end{cases}$$

Find CDF of X, E[X], Var[X]



$$F_X(x) = \begin{cases} 0 \ , & x < -\pi/2 \\ 1+\sin x/2 & -\pi/2 \leq x \leq \pi/2 \\ 1 \ , & x > \pi/2 \end{cases}$$

$$F_X(x) = \int_{-\pi/2}^{x} f_X(u)\,du = \frac{1}{2}\sin x\Big|_{-\pi/2}^{x} = \frac{1}{2}\Big(\sin x - \underbrace{\sin(-\pi/2)}_{-1}\Big)$$

$$= \frac{1+\sin x}{2}$$

$E[X] = \int_{-\pi/2}^{\pi/2} x \cdot \frac{1}{2}\cos x\,dx = 0.$ $\quad Var(X) = E[X^2]$

$$= \int_{-\pi/2}^{\pi/2} x^2 \cdot \frac{1}{2}\cos x\,dx = \frac{\pi^2}{4} - 2$$

 PAGE

DATE

★ Discrete - Continuous Joint Distributions

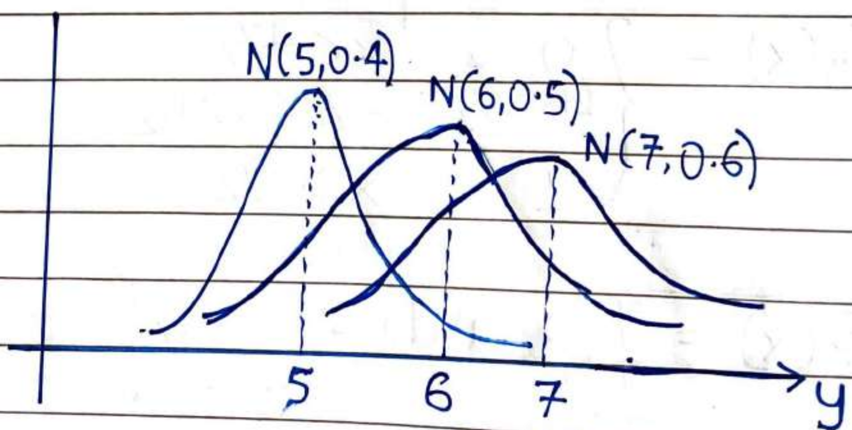
- (X,Y) : jointly distributed
- X : discrete with range T_X, PMF $p_X(x)$
- For each $x \in T_X$, we have a CRV Y_x, with density $f_{Y_x}(y)$
- Y_x : Y given $X=x$, i.e $(Y|X=x)$
- $f_{Y_x}(y)$: conditional density of Y given $X=x$, denoted $f_{Y|X=x}(y)$
- Marginal density of Y : $f_Y(y) = \sum_{x \in T_X} p_X(x) f_{Y|X=x}(y)$

Q. Let $X \sim \text{Unif}\{0,1,2\}$. Let $Y|X=0 \sim \text{Normal}(5, \overset{\sigma=}{0.4})$

$Y|X=1 \sim \text{Normal}(6, \overset{\sigma=}{0.5})$

$Y|X=2 \sim \text{Normal}(7, \overset{\sigma=}{0.6})$

- What is the marginal of Y?
- Suppose we observe Y to be around y_0. What can you say about X?



$$f_{Y|X=0}(y) = \frac{1}{\sqrt{2\pi} \times 0.4} e^{\frac{-(y-5)^2}{2(0.4)^2}}$$

$$f_Y(y) = \frac{1}{3}\cdot\frac{1}{\sqrt{2\pi}\times 0.4} e^{\frac{-(y-5)^2}{2\times(0.4)^2}} + \frac{1}{3}\cdot\frac{1}{\sqrt{2\pi}\times 0.5} e^{\frac{-(y-6)^2}{2(0.5)^2}} + \frac{1}{3}\cdot\frac{1}{\sqrt{2\pi}\times 0.6} e^{\frac{-(y-7)^2}{2\times(0.6)^2}}$$

"Not gaussian", "Mixture Gaussian"

★ Conditional Probability of Discrete given Continuous

Suppose X and Y are jointly distributed with $X \in T_X$ being discrete with PMF $p_X(x)$, conditional densities $f_{Y|X=x}(y)$ for $x \in T_X$. The conditional probability of X given $Y = y_0 \in \text{supp}(Y)$ is defined as :→

$$P(X=x \mid Y=y_0) = \frac{p_X(x)\, f_{Y|X=x}(y_0)}{f_Y(y_0)}$$

f_Y is marginal density of Y.

$\boxed{P(A|B)\,P(B) = P(B|A)\,P(A)}$

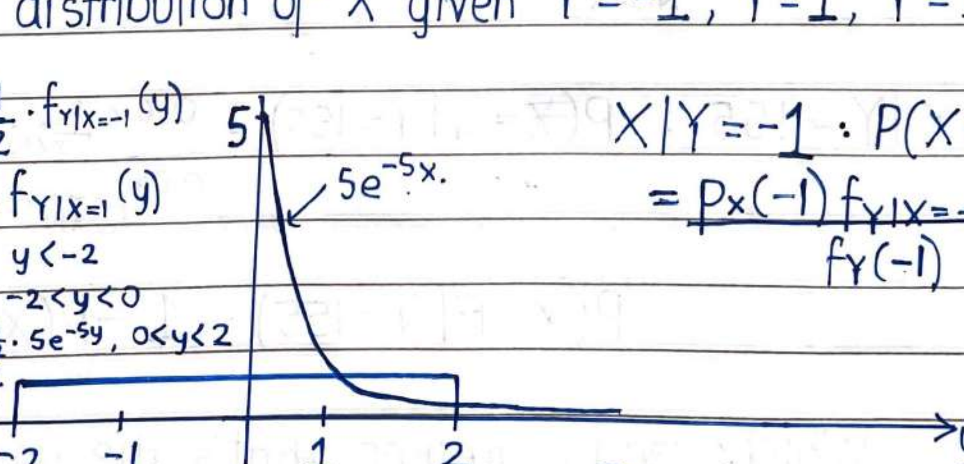
- Similar to Baye's Rule : $P(X=x|Y=y_0)\, f_Y(y_0) = p_{Y|X=x}(y_0)\, p_X(x)$

- X and Y independent? $f_{Y|X=x}$ is independent of x.
  $\therefore f_Y = f_{Y|X=x}$ and $P(X=x|Y=y_0) = p_X(x)$

Q. Let $X \sim \text{Unif}\{-1,1\}$. Let $Y|X=-1 \sim \text{Unif}[-2,2]$, $Y|X=1 \sim \text{Exp}(5)$
Find distribution of X given $Y=-1$, $Y=1$, $Y=3$.

$$f_Y(y) = \frac{1}{2}\cdot f_{Y|X=-1}(y) + \frac{1}{2} f_{Y|X=1}(y)$$

$$= \begin{cases} 0, & y<-2 \\ \frac{1}{2}\cdot\frac{1}{4}, & -2<y<0 \\ \frac{1}{2}\cdot\frac{1}{4}+\frac{1}{2}\cdot 5e^{-5y}, & 0<y<2 \\ \frac{1}{2}5e^{-5y}, & y>2 \end{cases}$$



$X|Y=-1$: $P(X=-1 \mid Y=-1)$

$$= \frac{p_X(-1)\, f_{Y|X=-1}(-1)}{f_Y(-1)} = \frac{\frac{1}{2}\cdot\frac{1}{4}}{\frac{1}{2}\cdot\frac{1}{4}} = \underline{\underline{1}}$$

$$P(X=+1 \mid Y=-1) = \frac{p_X(1)\, f_{Y|X=1}(-1)}{f_Y(-1)}$$

$$= \frac{1}{2}\cdot 0 \Big/ \frac{1}{2}\cdot\frac{1}{4} = \underline{\underline{0}}$$

classmate

DATE

$X|Y=1: \quad P(X=-1|Y=1) = \dfrac{\frac{1}{2}\cdot\frac{1}{4}}{\frac{1}{2}\cdot\frac{1}{4}+\frac{1}{2}\cdot 5e^{-5}}$

$P(X=1|Y=1) = \dfrac{\frac{1}{2}\cdot 5e^{-5}}{\frac{1}{2}\cdot\frac{1}{4}+\frac{1}{2}\cdot 5e^{-5}}$

$X|Y=3: \quad P(X=-1|Y=3) = 0$
$P(X=1|Y=3) = 1$

Q. 60% of adults in age 45-50 bracket in a country are male, 40% female. Height (in cm) of adult males in that age group & country is Normal (160, σ=10), and that of females is Normal (150, σ=5). Random person's height is 155 cm. Is that person more likely to be male or female?

$X \sim \{\overset{0.6}{M}, \overset{0.4}{F}\}$ $\quad Y|X=M \sim N(160, \sigma=10)$
$Y|X=F \sim N(150, \sigma=5)$

$$f_{Y|X=M}(y) = \frac{1}{\sqrt{2\pi}\times 10} e^{\frac{-(y-160)^2}{2(10)^2}}$$

$$f_{Y|X=F}(y) = \frac{1}{\sqrt{2\pi}\times 5} e^{\frac{-(y-150)^2}{2(5)^2}}$$

i.e $(155-160)^2$

$$X|Y=155: \quad P(X=M|Y=155) = \frac{0.6\times\frac{1}{\sqrt{2\pi}\times 10} e^{\frac{-5^2}{2\cdot 10^2}}}{0.6\times\frac{1}{\sqrt{2\pi}\times 10} e^{\frac{-5^2}{2\cdot 10^2}} + 0.4\times\frac{1}{\sqrt{2\pi}\cdot 5} e^{\frac{-5^2}{2\cdot 5^2}}}$$

$$P(X=F|Y=155) = 1 - P(X=M|Y=155)$$

Whichever is greater, that's the answer.

DATE

Q. Let $Y = X + Z$, $X \sim \text{Unif}\{-3, -1, 1, 3\}$ and $Z \sim \text{Normal}(0, \sigma^2)$ are independent. Distribution of Y? Of $(X \mid Y = 0.5)$

$f_{Y|X=-3}(y) = ?$

$Y|_{X=-3} \leftrightarrow (-3+Z) \sim N(-3, \sigma^2)$

$Y|_{X=-1} \leftrightarrow (-1+Z) \sim N(-1, \sigma^2)$

$Y|_{X=1} \leftrightarrow N(1, \sigma^2)$, $Y|_{X=3} \leftrightarrow N(3, \sigma^2)$

The rest is the same as before.