21.11.2022

YZV 231E

Probability Theory & Stats

Week 9

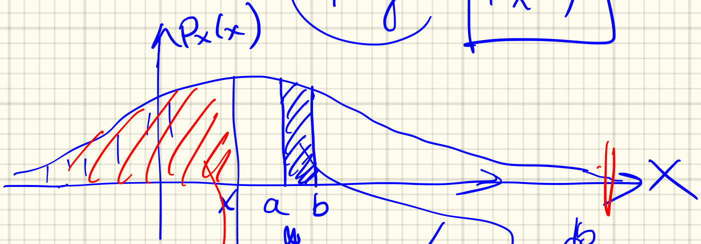
Gü.

Recap: Continuous r.v.s

X r.v. $\rightarrow$ pdf $\boxed{p_X(x)}$

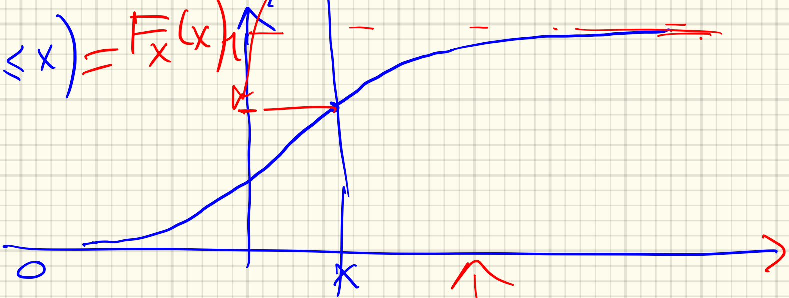
$p_X(x) \geq 0$

$\int_{-\infty}^{\infty} p_X(x)\,dx = 1$



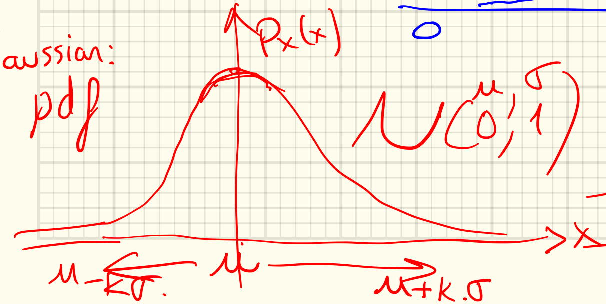
$P(a < X < b) = \int_a^b p_X(x)\,dx$

CDF: $P(X \leq x) = F_X(x)$

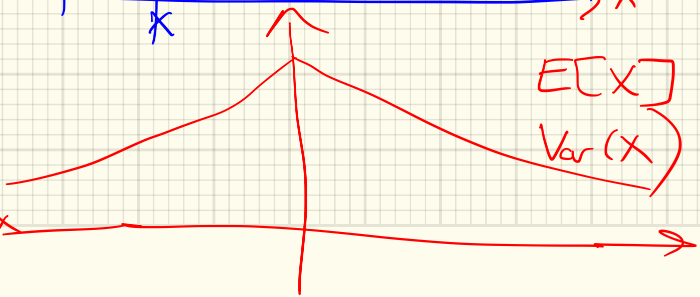


Gaussian:

pdf



$\mathcal{N}(\overset{\mu}{0}, \overset{\sigma}{1})$



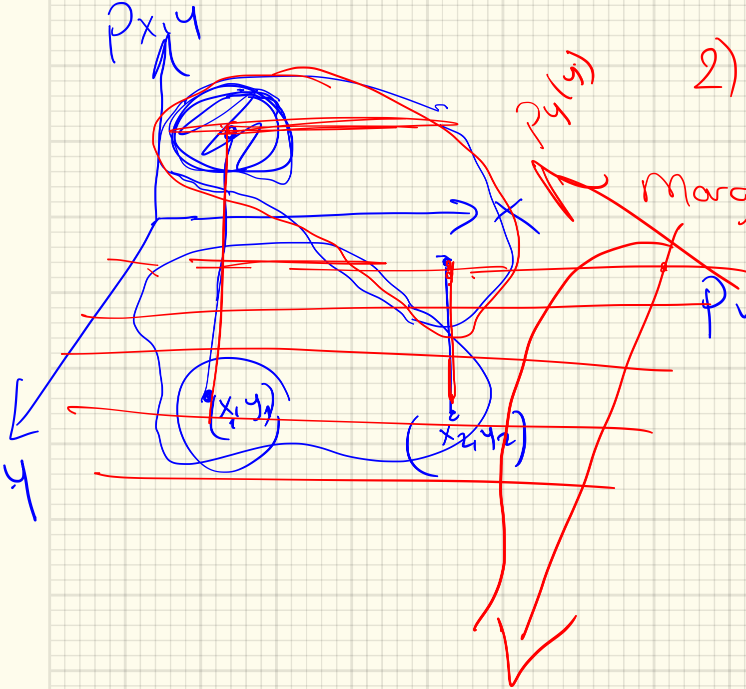
$E[X]$

$Var(X)$

Joint pdfs : X, Y r.v.s

1) $\boxed{p_{X,Y}(x,y)} \geq 0$

2) $\iint p_{X,Y}(x,y)\,dx\,dy = 1$

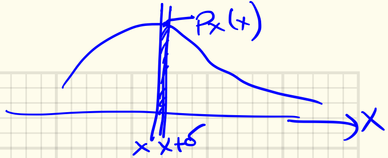


Marginal pdf :

$$p_Y(y) = \int_{-\infty}^{\infty} p_{X,Y}(x,y)\,dx$$

Independence: X, Y are independent : $p_{X,Y}(x,y) = p_X(x) \cdot p_Y(y)$

Conditional pdfs:



Recall: $P(x < X < x+\delta) \approx p_X(x)\cdot\delta$

$$P(x < X < x+\delta \mid \underbrace{Y=y}_{\text{Given } Y\approx y}) \approx \boxed{p_{X|Y}(x|y)}\cdot\delta$$

← conditional universe

1) $p_{X|Y}(x|y) \geq 0$

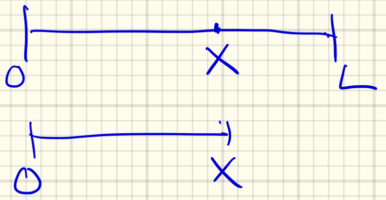
2) $\int_{-\infty}^{\infty} p_{X|Y}(x|y)\,dx = 1$

$$\Rightarrow \boxed{p_{X|Y}(x|y) = \frac{p_{X,Y}(x,y)}{p_Y(y)+\varepsilon}}$$

if $p_Y(y) > 0$

Ex: Start w/ a stick of certain length L;

we break it at a random location X;



$$X \sim \text{Uniform}[0, L]$$

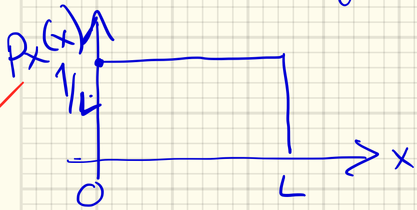
we break it again at a random location Y.

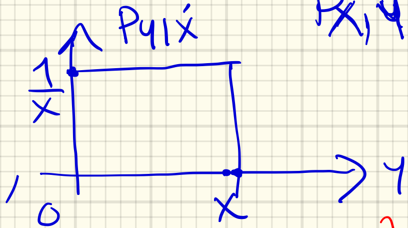
$$Y|X \sim \text{Uniform}[0, X]$$

Q. Joint pdf of X & Y?

$$p_{X,Y}(x,y) = p_{Y|X}(y|x) \cdot p_X(x)$$

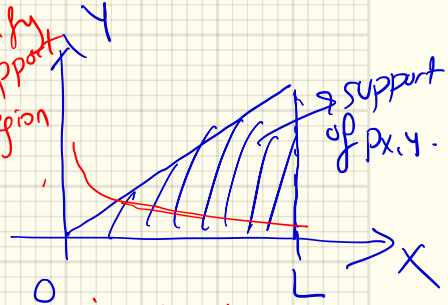
$$\left(= p_{X|Y}(x|y) \cdot p_Y(y) \right)$$





$$\boxed{p_{X,Y}(x,y) = \frac{1}{L} \cdot \frac{1}{x}, \quad x \in [0, L], \; y \in [0, x]}$$

specify support region



$x = 0$ $\quad p_{X,Y} \nearrow \infty$

$x = L$ $\quad p_{X,Y} = \frac{1}{L^2}$

exercise: imagine in 3D.

$E[Y \mid X=x] = ? = \int_0^x y \cdot p_{Y|X}(y \mid X=x)\, dy$

$= \int_0^x y \cdot \frac{1}{x}\, dy = \frac{1}{x} \frac{y^2}{2} \Big|_0^x = \frac{x}{2}$ ✓

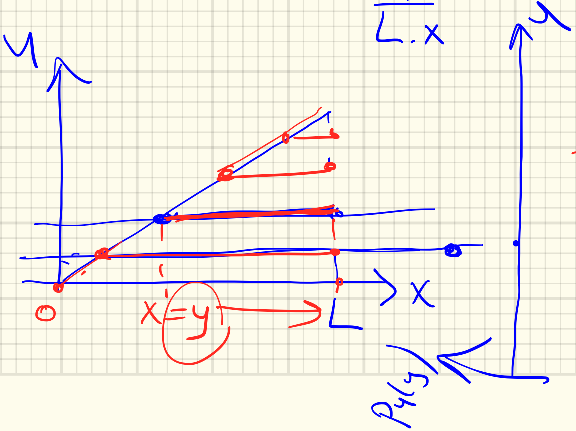
Intuitive ✓

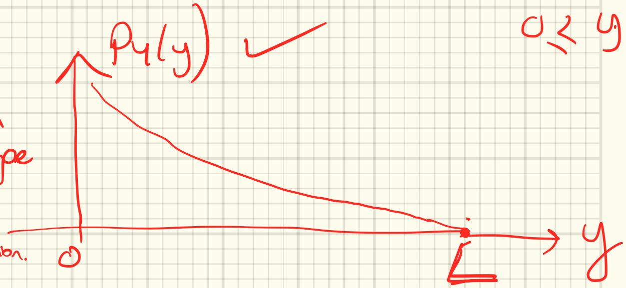
cm of the cond. pdf $= \frac{x}{2}$ ✓

Q. Marginal density of Y : $p_Y(y)$?

$p_Y(y) = \int \underbrace{p_{X,Y}(x,y)}_{\frac{1}{L.x}}\, dx = \int_y^L \frac{1}{Lx}\, dx = \frac{1}{L} . \ln x \Big|_y^L = \underbrace{\frac{1}{L} \ln \frac{L}{y}}_{0 \le y \le L.}$

pay attention to range of integration.





$E[Y] = ? = \int y\, p_Y(y)\, dy = \int_0^L \frac{1}{L} \ln \frac{L}{y}\, dy = \frac{L}{4}$

IBP : $-\int y . \ln y\, dy$

Continuous Bayes Rule : Make Inferences

Unknown / unobserved signal information X → Random phenomenon → Y : observation

prior / belief $p_X(x)$

Measurement Device

we observe r.v. Y

$Y|X$: likelihood.

$p_{Y|X}$ conditional pdf.

Target

$p_{X|Y}$: posterior density

making inferences about X.

— Recall:

Discrete case we worked on the ex

$X = 1/0$: airplane present or not

$Y = 1/0$: radar fired or not.

$$\underbrace{p_{X|Y}(x|y)}_{\text{posterior density}} = \frac{\overbrace{p_{Y|X}(y|x)}^{\text{likelihood}} \cdot \overbrace{p_X(x)}^{\text{prior}}}{\underbrace{p_Y(y)}_{\text{Evidence}}}$$

$$\underbrace{p_{X|Y}(x|y)}_{\text{unobserved.}}$$

Evidence : $p_Y(y) = \int_x p_{Y|X}(y|x) \cdot p_X(x) \, dx$

Continuous Inference: Standard example in signal/data processing/communications

X : some unobserved signal : "prior" $p_X(x)$

Y : observed "noisy" version of the signal X : $p_{Y|X}(y|x)$: model of the noise

$p_{X|Y}(x|y)$

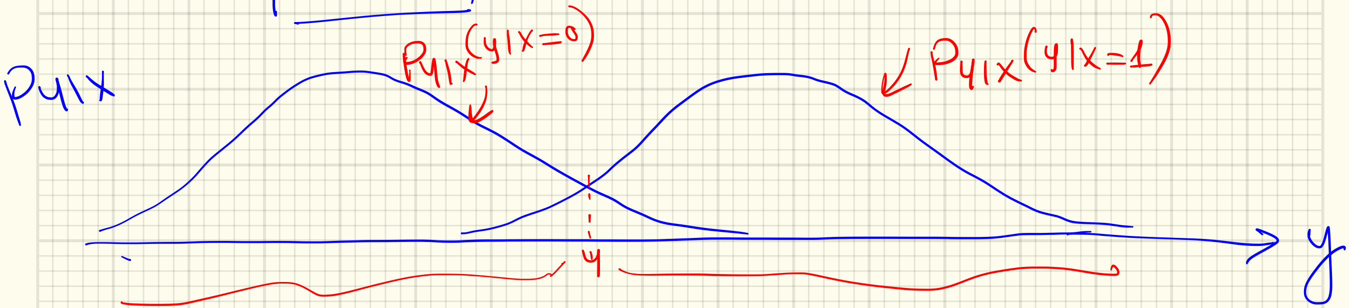
Ex: Discrete $X = 0, 1 \rightarrow$ xmitted

We measure Continuous $\underbrace{Y = X + W}_{\text{model of the measurement device}}$

$X \xrightarrow{\text{xmitter}} \boxed{\text{Commun. channel}} \rightarrow Y = X + W$ Rcvr.

Say $W \sim$ Gaussian distributed noise.

$X \longrightarrow \boxed{Y|X} \longrightarrow Y$



$P_{X,Y}(x,y)$ for X: discrete r.v. ; Y: continuous r.v.

$$P_{X,Y}(\underbrace{X=x}_{\text{discrete}}, \underbrace{y \le Y \le y+\delta}_{\text{continuous}}) = \underbrace{P_X(X=x)}_{\text{discrete marginal pmf}} \cdot \underbrace{P(y \le Y \le y+\delta \mid X=x)}_{\text{conditional continuous pdf}}$$

$$= \underbrace{p(y \le Y \le y+\delta)}_{\text{continuous marginal}} \cdot \underbrace{p(X=x \mid y \le Y < y+\delta)}_{\text{discrete conditional}}$$

Posterior:

$$P_{X|Y}(x|y) = \frac{P_{Y|X}(y|x)\, P_X(x)}{P_Y(y)}.$$

Ex: If Continuous X, Discrete Y

X (light intensity) $\longrightarrow$ $\boxed{\text{Measuring Device}}$ $\longrightarrow$ Y ; photon count.

Target: We try to infer $X \mid Y$.

→ Same Bayes formula:

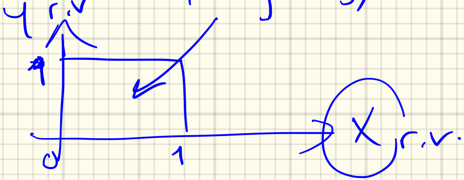
$$\underbrace{P_{X|Y}(x|y)}_{\text{posterior}} = \frac{P_{Y|X}(y|x)\, P_X(x)}{\int P_{Y|X}(y|x)\, P_X(x)\, dx}$$

$$\int P_{Y|X}(y|x)\, P_X(x)\, dx = P_Y(y)$$

★ How you model these probability density functions (whether x's, y's are discrete / conti.)
or pmf s
priors & conditionals
to infer $\boxed{\text{posterior density function}}$

Distribution of Transformed R.V.'s $\equiv$ Derived Distributions.

Given $p_{X,Y}(x,y)$



$g(X,Y) = \frac{Y}{X} = Z$ r.v.: ratio of X & Y.

X: r.v.
Y: r.v.

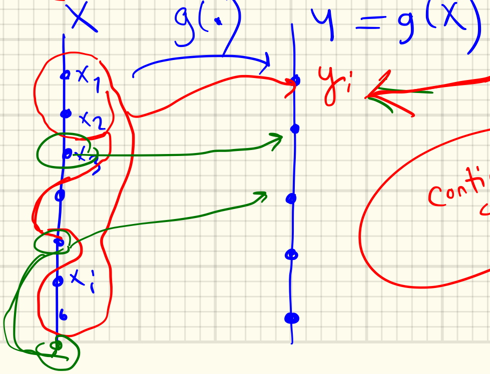
→ a new r.v. you derived from X & Y.

pdf of Z?

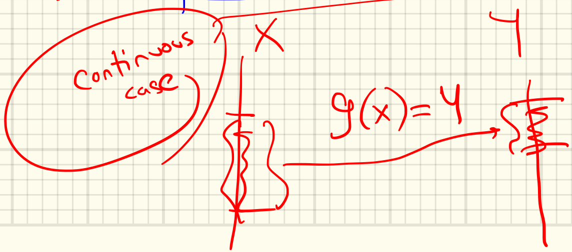
ex: $E[\underbrace{g(X,Y)}_{Z}] = \iint g(x,y)\, p_{X,Y}(x,y)\, dx\, dy$ ✓ Don't need the xformed pdf of Z

→ Now we want the distribution of $Z = g(X,Y)$.

Recall: Discrete r.v.'s case



$$\boxed{p_Y(y_i) = \sum_{x_i \in g^{-1}(y_i)} p_X(x_i)}$$

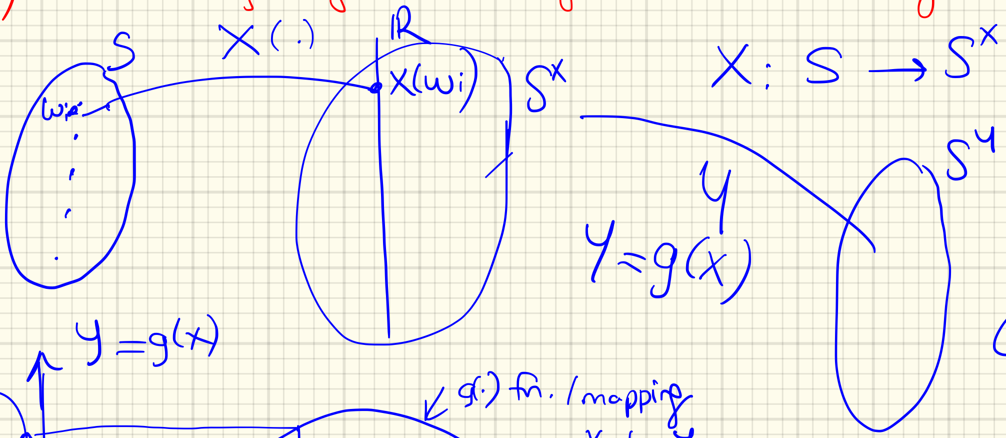


2 ways

1) Change of variables

2) CDF

1) Use change of variables formula to transform r.v.s.

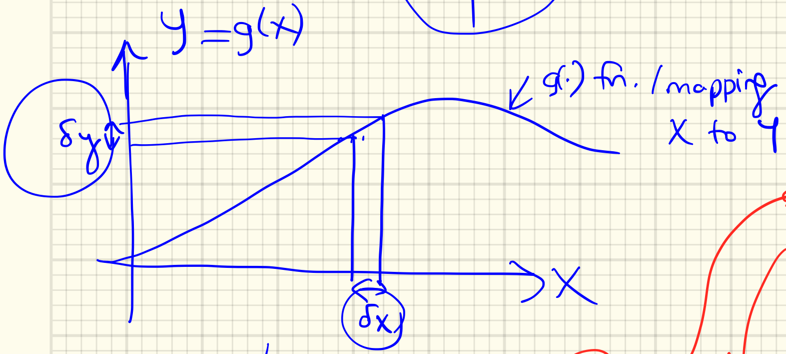


$$X : S \rightarrow S^X \subseteq \mathbb{R}.$$

Given r.v. X & $p_X(x)$:

Q: $p_Y(y) = ?$

$y = g(x)$



$$\boxed{p_Y(y) = p_X(x) \left| \frac{dx}{dy} \right|} \rightarrow \text{Jacobian}$$

$$\left(p_X(x) = p_Y(y) \left| \frac{dy}{dx} \right| \right)$$

$$\underbrace{p_X(x) \cdot \delta_x}_{\text{total probability}} = \underbrace{p_Y(y) \cdot \delta_y}_{\text{Probability}}$$

$x < X < x + \delta_x$

$$\boxed{p_Y(y) = p_X(\underbrace{g^{-1}(y)}_{x}) \cdot \left| \frac{dg^{-1}(y)}{dy} \right|}$$

for g(.) is one-to-one (invertible)

$$(x_1, \dots, x_n) \longrightarrow (y_1, \dots, y_n)$$: Jacobian matrix: matrix of 1st order partial derivatives

$$J = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_2}{\partial x_n} \\ \vdots & & & \\ \frac{\partial y_m}{\partial x_1} & \cdots & \cdots & \frac{\partial y_m}{\partial x_n} \end{bmatrix}$$

Jacobian Matrix.

$$X \to Y = g(x)$$

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

$$P_Y(y) = P_X(\underset{g^{-1}(y)}{x}) \underbrace{\left| \frac{dx}{dy} \right|}_{\left| \frac{1}{\frac{dy}{dx}} \right|} = \left| \frac{d\, g^{-1}(y)}{dy} \right|$$

Ex: Affine transformation of r.v.s.

(Linear + offset)

$$Y = aX + b \qquad a \in \mathbb{R}, a \neq 0, \; b \in \mathbb{R}$$

Given X & $p_X(x)$; $p_Y(y) = ?$

$y = g(x) = ax + b$

$x = g^{-1}(y) = \frac{y-b}{a}$

$\frac{d g^{-1}(y)}{dy} = \frac{1}{a}$

$$p_Y(y) = p_X\left(\frac{y-b}{a}\right) \underbrace{\left| \frac{d \overbrace{g^{-1}(y)}^{x}}{dy} \right|}_{= \frac{1}{|a|}}$$

$$\boxed{p_Y(y) = \frac{1}{|a|} p_X\left(\frac{y-b}{a}\right)}$$

ex: $Y = aX + b$, $a \neq 0$, $X \sim \mathcal{N}(0,1) = \frac{1}{\sqrt{2\pi}} e^{-y^2/2}$ $\rightarrow Y = ?$

$S_X = (-\infty, \infty) \rightarrow S_Y = (-\infty, \infty)$

$$p_Y(y) = \frac{1}{|a|} p_X\left(\frac{y-b}{a}\right) = \underbrace{\frac{1}{|a|} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{y-b}{a}\right)^2}}_{} \sim \mathcal{N}(b, a^2) \checkmark$$

$\rightarrow$

★ Affine-transformed Gaussian r.v.s are still Gaussian r.v.'s w/ a new mean & variance.

Ex: $X \sim \mathcal{N}(0,1) \Rightarrow Y = e^X$, $\quad y = g(x) = e^x$

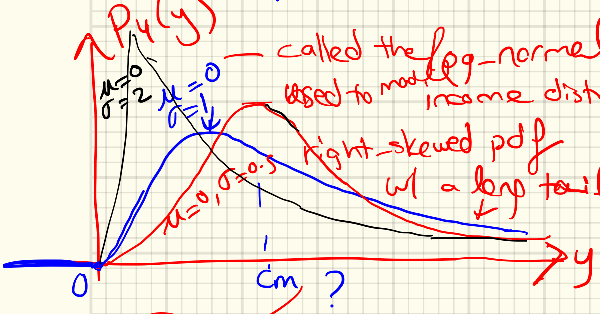
$S_x = (-\infty, \infty)$; $S_y = (0, \infty)$

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

$\dfrac{dg^{-1}}{dy} = \dfrac{1}{y}$

$$P_Y(y) = \begin{cases} P_X(\ln y) \cdot \dfrac{1}{y} & , y<0 \\ 0 & , y \geq 0 \end{cases}$$

$$\boxed{P_Y(y) = \begin{cases} \dfrac{1}{\sqrt{2\pi}} \dfrac{1}{y} e^{-\frac{1}{2}(\ln y)^2} & , y>0 \\ 0 & , y \leq 0 \end{cases}}$$



— called the log-normal density

— used to model income distrib; stock prices ... → b/c these cannot be -ve!

right-skewed pdf w/ a long tail

In general: $\boxed{P_Y(y) = \dfrac{1}{y\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln y - \mu}{\sigma}\right)^2}}$: μ : location, σ : scale. parameters

* Note: Always determine the support of pdfs

$$p_X(x) \longrightarrow p_Y(y)$$

Given $S_X = \underline{\qquad\qquad}$ $\longrightarrow$ $S_Y = ?$

exercise: start w/ $Y \sim \text{logNormal}(\overset{\mu}{0}, \overset{\sigma}{1})$

$\hookrightarrow X = g(Y) = \ln Y$ $\rightarrow$ then X is $\mathcal{N}(0,1)$

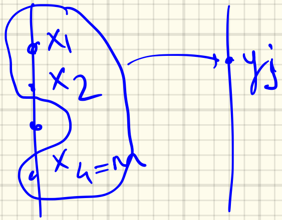
* If the transformation $g(\cdot)$ is $\boxed{\text{many-to-one}}$ mapping

$$y = g(x) \quad ; \quad \text{e.g. } g(x) = x^2 = y \Rightarrow g^{-1}(y) = x.$$

$$x = \mp\sqrt{y}$$

$$\{x : g(x) = y\}$$

$$P_Y(y) = \sum_{x \in g^{-1}(y)} P_X(x) \left| \frac{1}{\frac{dg(x)=y}{dx}} \right| \rightarrow \left|\frac{dx}{dy}\right| = \left|\frac{dg^{-1}(y)}{dy}\right|$$



$$x_i = g_i^{-1}(y_j) \quad , \quad \text{for } i = 1, \dots M$$

$$\boxed{P_Y(y) = \sum_{i=1}^{M} P_X\left(g_i^{-1}(y)\right) \left| \frac{dg_i^{-1}(y)}{dy} \right|}$$

again sum for all M elements in $g^{-1}(y)$.

Ex :

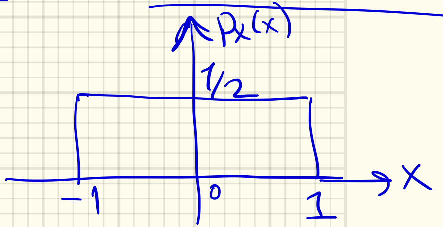
$Y = X^2 = g(x)$, $X \sim U[-1, 1]$

$S_x = [-1, 1] \longrightarrow S_y = [0, 1]$

Defined the support for y

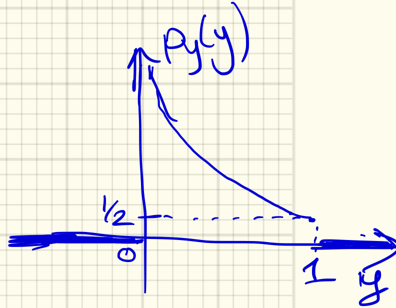
$x_1 = \sqrt{y} = g^{-1} \longrightarrow \frac{dg^{-1}}{dy} = \frac{1}{2\sqrt{y}}$

$x_2 = -\sqrt{y} = g^{-1} \longrightarrow \frac{dg^{-1}}{dy} = -\frac{1}{2\sqrt{y}}$



$$p_Y(y) = p_X(\sqrt{y}) \cdot \left|\frac{1}{2\sqrt{y}}\right| + p_X(-\sqrt{y}) \left|\frac{1}{-2\sqrt{y}}\right|$$

$$= \frac{1}{2}\frac{1}{2\sqrt{y}} + \frac{1}{2}\frac{1}{2\sqrt{y}} = \frac{1}{2\sqrt{y}}, \quad y \in (0, 1]$$



exercise : $Y = X^2$, $X \sim \mathcal{N}(0, 1)$

Derive $p_Y(y)$.

[S Kay Example 10.7]

(2) 2nd way to transform r.v.s X to get the Xformed pdf:

CDF Approach: $X \xrightarrow{g} Y$.

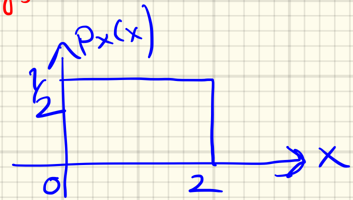
2 step procedure:

1) Get CDF of Y: $F_Y(y) = P(Y \leq y)$ (using CDF/pdf of X)

2) Differentiate to get $p_Y(y) = \dfrac{dF_Y(y)}{dy}$

Ex: X: Uniform on $[0,2]$

$S_X = [0,2]$



$S_X = [0,2]$

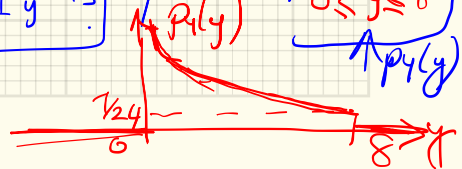
$S_Y = [0,8]$

$Y = X^3$, Find $p_Y(y)$.

1st Way: $p_Y(y) = p_X(x)\left|\dfrac{dg^{-1}(y)}{dy}\right| = p_X(y^{1/3})\left|\dfrac{1}{3\, y^{2/3}}\right| = \dfrac{y^{-2/3}}{6}$, $0 \leq y \leq 8$

$\frac{1}{2}$

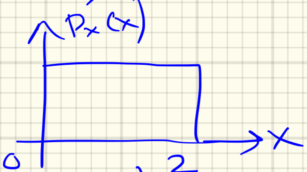
$g^{-1}(y) = \boxed{x = y^{1/3}} \xrightarrow{\frac{d}{dy}} \boxed{\frac{1}{3} y^{-2/3}}$



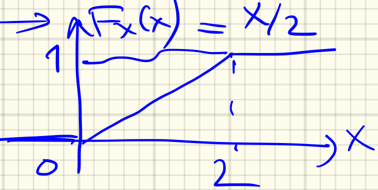
2nd way: (i) $F_Y(y) = P(Y \leq y) = P(X^3 \leq y)$

$$= P(X \leq y^{1/3}) = F_X(y^{1/3})$$

$$F_Y(y) = \begin{cases} \frac{1}{2} \cdot y^{1/3}, & 0 \leq y \leq 8 \\ 0, & y < 0 \\ 1, & y > 8 \end{cases}$$



cdf of X. $\rightarrow$



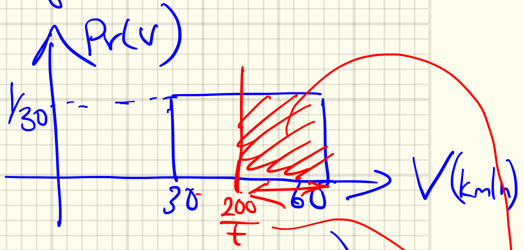
(ii) $p_Y(y) = \frac{d}{dy}\left(\frac{1}{2} y^{1/3}\right) = \begin{cases} \frac{1}{6y^{2/3}}, & y \in (0,8) \\ 0, & \text{o/w.} \end{cases}$

Ex: Ece driving from Istanbul to Edirne (dist $= 200$ km).

Her speed is uniformly distrib. btw $(30, 60)$ km/h.

What is the distribution of the duration of the trip?

$$V \sim U[30, 60] \to p_V(v)$$



Duration of the trip

$$T(V) = \frac{200}{V} \to p_T(t) = ?$$

Use CDF way:

i) $$F_T(t) = P(T \le t) = P\left(\frac{200}{V} \le t\right)$$

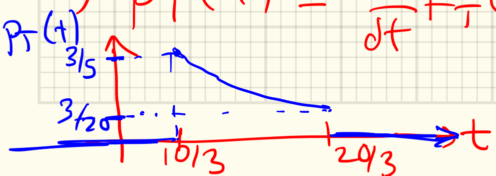
$$= P\left(V \ge \frac{200}{t}\right)$$

$$\to F_T(t) = \frac{1}{30} \cdot \left(60 - \frac{200}{t}\right)$$

ii) $$p_T(t) = \frac{d}{dt} F_T(t) = \frac{200}{30} \cdot \frac{1}{t^2}$$

$$\frac{10}{3} \le t \le \frac{20}{3}$$

$200/60$ $= 200/30$



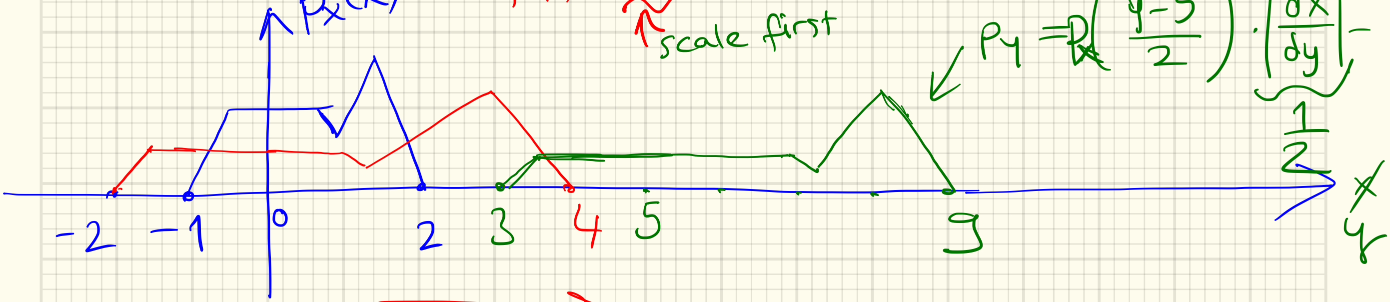
exercise:

$$E[T] = ? \quad Var(T) = ?$$

Ex: $y = 2x + 5$

$p_x(x) \longrightarrow p_x(y/2)$ — scale first — then shift $\longrightarrow$

$$p_y = p_x\left(\frac{y-5}{2}\right) \cdot \underbrace{\left|\frac{dx}{dy}\right|}_{\frac{1}{2}}$$



expanded my axis by 2

Distribution of $X+Y$: Transformation of multiple r.v.s.

$W = X+Y$; $\left(W = g(X,Y) \right)$ Given: X, Y pdf's are known.

X & Y are independent. $\longrightarrow$ find $p_W(w)$.

3) Use Conditioning to find $p_W(w)$: General way to xform $g(X,Y) \to W$

i) Fix $X=x$, let $W|X=x = g(\underset{\uparrow \text{const.}}{x}, Y) = g_x(Y)$: function of Y. x is fixed.

ii) Find $P_{W|X}(w|x)$: using transformation of $Y \longrightarrow W = g_x(Y)$.

iii) Uncondition to find $p_W(w) = \int_{-\infty}^{\infty} \underbrace{p_{W|X}(w|x) \cdot p_X(x)}_{p_{W,X}(w,x)} \, dx$

e.g. $W = \underset{\uparrow x:\text{const.}}{X} + Y = g_x(Y)$, X & Y are indep. $\longrightarrow p_{Y|X} = p_Y$

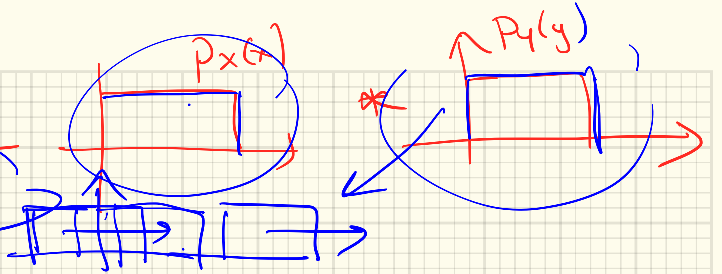
$$p_W(w) = \int_{-\infty}^{\infty} \underbrace{p_{W|X}(w|x)}_{p_{Y|X}(\underbrace{w-x}_{y}) = p_Y(w-x)} p_X(x) \, dx = \boxed{\int_{-\infty}^{\infty} p_Y(w-x) \, p_X(x) \, dx.}$$

$$\boxed{p_W(w) \triangleq p_X(x) * p_Y(y)}$$ convolution operation.

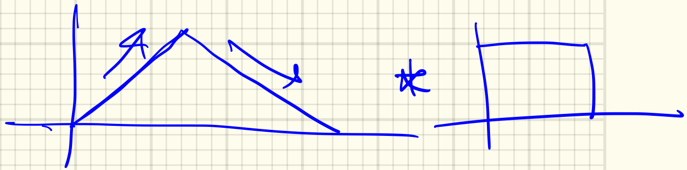
Convolution Operation:

$W = X + Y$

X, Y independent



$$P_W(w) = \int P_X(z)\, P_Y(w-z)\, dz = P_X * P_Y.$$



$$W = X_1 + X_2 + \cdots + X_n$$

identically distr

independent.

$n \to \infty$.

Gaussian.