Probability and Random Process

Aimin Tang

The University of Michigan- Shanghai Jiao Tong University Joint Institute
Shanghai Jiao Tong University

Sep. 27 2020

Ve501 2020-2021 Fall

Based on Lecture Notes by Prof. Chong Han

Outline

- 2. Random Variables
  - Introduction to Random Variables
  - PMF and Discrete Random Variables
  - PDF and Continuous Random Variables
  - Gaussian CDF
  - Conditional Probability
  - Function of a RV
  - Expectation of a RV
  - Transform Methods and Probability Generating Function

Function of a RV

Function of a RV

- The problem:
  - Given $f_X(x)$ and $Y = G(X)$,
  - find $f_Y(y)$

- Example application: X is voltage, Y is associated power through a 1Ω resistor.

$$X \sim N(0, \sigma^2)$$

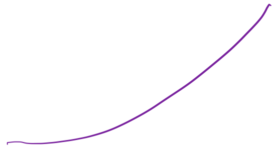
$$Y \sim \text{Chi Square}$$

$$\sqrt{\mathrm{Y}} \sim \text{Rayleigh}$$

CDF for Function of a RV

- Find the corresponding set of X
  $\Pr(Y \in A) = \Pr(X \in G^{-1}(A)) = \Pr(X \in \{x : G(x) \in A\})$
  inverse image
- Determine what are the possible values of Y, i.e., what kind of r.v. is Y
  - If X is discrete, then Y is discrete
  - If X is continuous, then Y can be discrete, continuous or mixed

Example



单增.

- $Y = g(X)$ where $g(x) = 2e^{3x}$, is a function of random variable X. Find the CDF of Y in terms of CDF of X

$\frac{1}{3}\ln(\frac{y}{2})$

$$F_Y(y) = \begin{cases} F_X(g^{-1}(y)) & y>0 \\ 0 & y\le 0 \end{cases}$$

PMF for discrete X

if discrete.

The pmf of Y is

$$p_Y(y) = \mathrm{P_r}(Y = y) = \mathrm{P_r}\left(X \in g^{-1}(\{y\})\right)$$

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

PDF for continuous X

1. Find CDF $F_Y(y)$ and differentiate
   - CDF works for all r.v.'s

     无论分布 discrete / continuous
     都有 CDF

2. The method of differentials

Example

- X ~ Uniform[−1, 1], and Y = g(X) where $g(x) = 2e^{3x}$

  Find the pdf of Y in terms of pdf of X.

$$f_Y(y) = \begin{cases} \frac{d}{dy} F_X\left(\frac{1}{3}\ln\frac{y}{2}\right) \\ 0 \quad \text{otherwise} \end{cases} = \begin{cases} \frac{1}{3}\cdot\frac{2}{y}\cdot\frac{1}{2} f_X\left(\frac{1}{3}\ln\frac{y}{2}\right) \\ 0 \end{cases}$$

$$\frac{1}{3}\ln\frac{y}{2} \in [-1, 1]$$

$$-3 \le \ln\frac{y}{2} \le 3$$

$$2e^{-3} \le y \le 2e^{3}$$

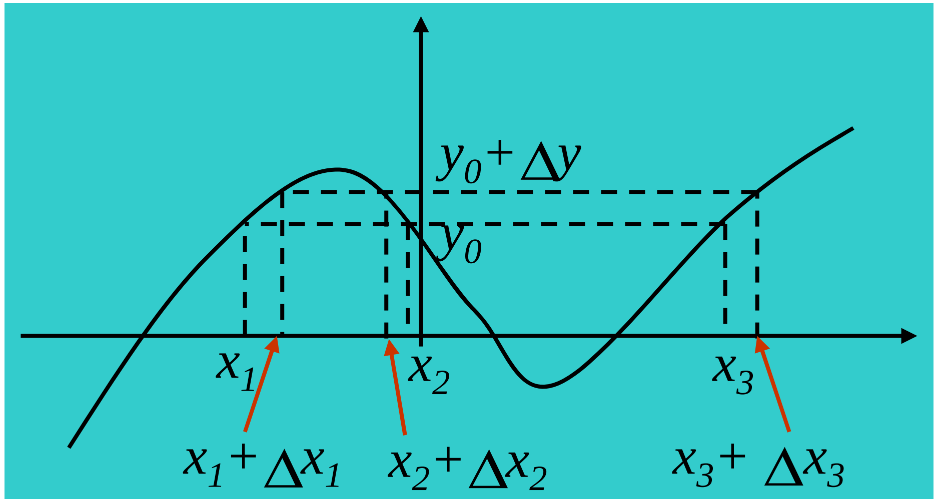
$$= \begin{cases} \frac{1}{6y} & y \in [2e^{-3}, 2e^{3}] \\ 0 & \text{otherwise} \end{cases}$$

The method of differentials - I

- Start with a differential interval on the Y-axis.

$$y_0 \le Y \le y_0 + \Delta y$$

- Identify all values of X that map into that differential Y interval.

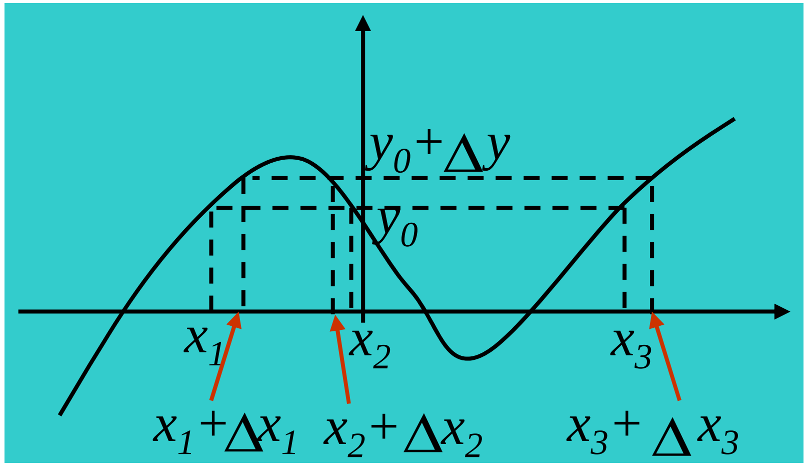


x_1, x_2, and x_3 are solutions to $Y=G(X)$

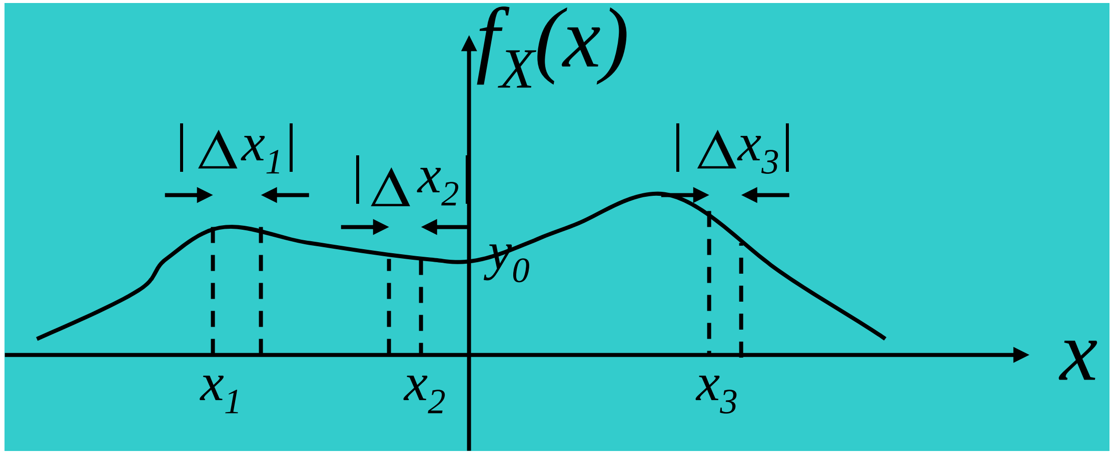
The method of differentials - II

$$P(y_0 \le Y \le y_0 + \Delta y) = P(x_1 \le X \le x_1 + \Delta x_1 \\ \cup x_2 \le X \le x_2 + \Delta x_2 \\ \cup x_3 \le X \le x_3 + \Delta x_3)$$

$$= P(x_1 \le X \le x_1 + \Delta x_1) + P(x_2 \le X \le x_2 + \Delta x_2) \\ + P(x_3 \le X \le x_3 + \Delta x_3)$$



The method of differentials - III



- Assuming the PDF is smooth enough, and Δx is small enough,

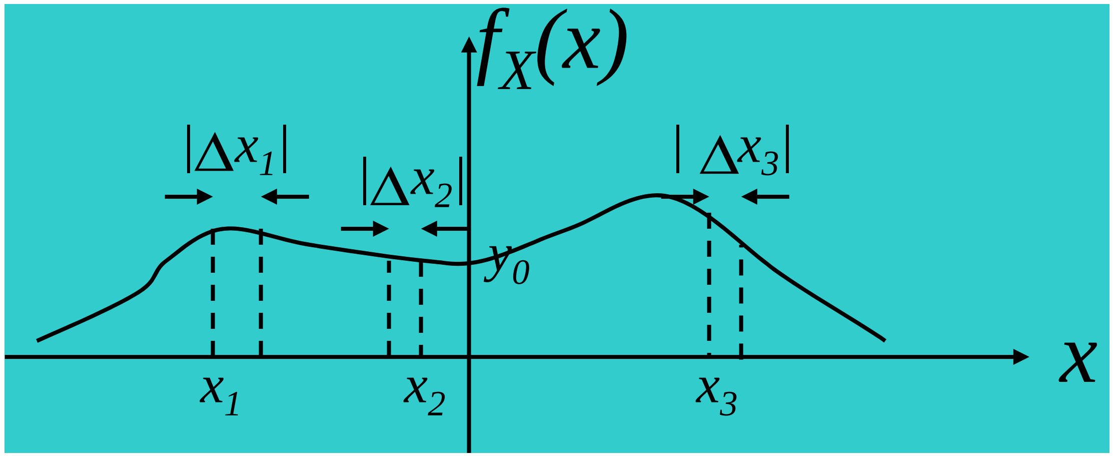
$$P(x_i \le X \le x_i + \Delta x_i) \approx f(x_i)\Delta x_i$$

The method of differentials - IV

- Δx_i is related to Δy through the slope of the function:

$$P(y_0 \le Y \le y_0 + \Delta y) \approx f_Y(y_0)\Delta y \approx \sum_{i=1}^{3} f_X(x_i)\left|\Delta x_i\right|$$

$$\approx \sum_{i=1}^{3} f_X(x_i)\frac{\left|\Delta y\right|}{\left|\frac{dy}{dx}\right|_{x=x_i}}$$



$\frac{\Delta y}{dy} = \frac{\Delta x_i}{dx_i}$

The method of differentials - V

Now,

$$f_Y(y_0)\Delta y \approx \sum_{i=1}^{3} f_X(x_i)\frac{|\Delta y|}{\left|\frac{dy}{dx}\right|_{x=x_i}}$$

As $\Delta y \to 0$, "$\approx$" becomes "$=$" and the result is:

$$f_Y(y_0) = \sum_{i=1}^{3} \frac{f_X(x_i)}{\left|\frac{dy}{dx}\right|_{x=x_i}}$$

General Rule

Given a function $Y = G(X)$ with continuous and smooth variation (derivative exists) and a continuous RV X,

$$f_Y(y) = \sum_{i=1}^{n} \frac{f_X(x_i)}{\left|\frac{dy}{dx}\right|_{x=x_i}}$$

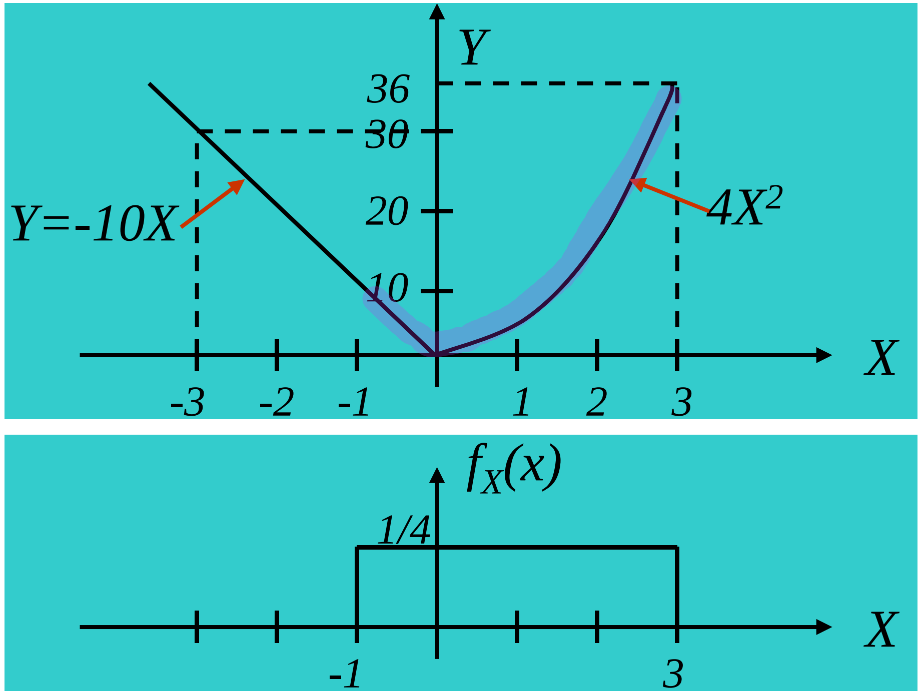
Where n is the number of solutions to $Y = G(X)$.

REMEMBER DO NOT APPLY TO

1. Flat parts of $Y=G(X)$ otherwise derivative $=0$.
2. Delta functions in $f_X(x)$

Function of a RV Examples

Ex. 1:



$$Y = G(x) = \begin{cases} -10x & x \le 0 \\ 4x^2 & x > 0 \end{cases}$$

$$f_Y(y) = \frac{\sum f_X(x)}{\left|\frac{dy}{dx}\right|}$$

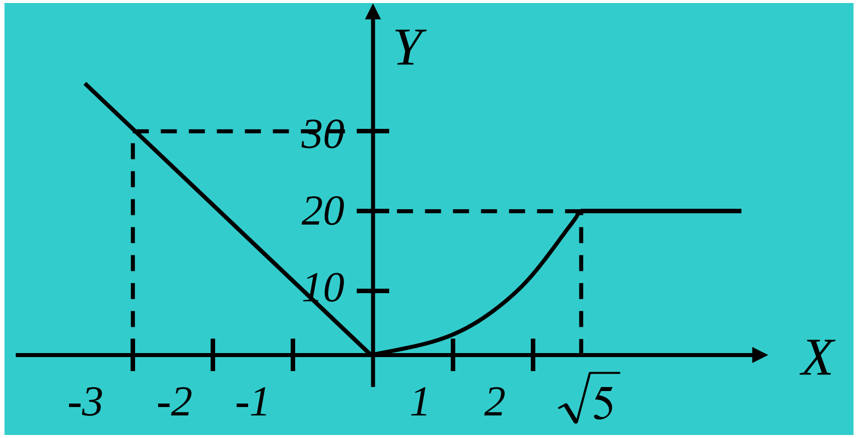
$$= \begin{cases} \cancel{\frac{\frac{1}{4}}{|-10|} = \frac{1}{40}} & -1 < x \le 0 \rightarrow 0 < y \le 10 \\ \cancel{\frac{\frac{1}{4}}{8x} = \frac{1}{32x}} & 0 < x \le 3 \rightarrow 0 < y \le 36. \end{cases}$$

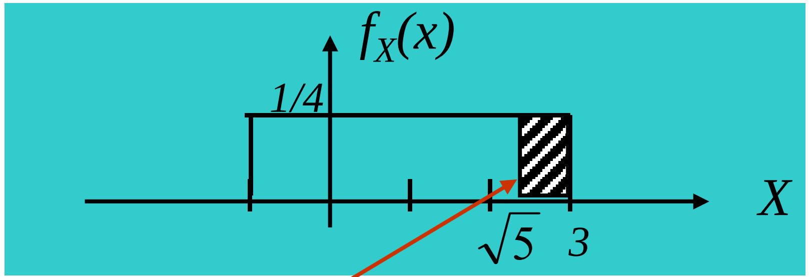
$$f_Y(y) = \begin{cases} \frac{1}{40} + \frac{1}{32\sqrt{\frac{y}{4}}} & 0 < y \le 10 \\ \frac{1}{32\sqrt{\frac{y}{4}}} & 10 < y \le 36 \\ 0 & \text{otherwise.} \end{cases}$$

Observe that $f_Y(y) = 0$ for $y > 36$ and $y < 0$ because no probability mass maps to these regions.

Example – 2

Same as Ex 1 but function has a flat part:





Shaded area= $\dfrac{3-\sqrt{5}}{4}$

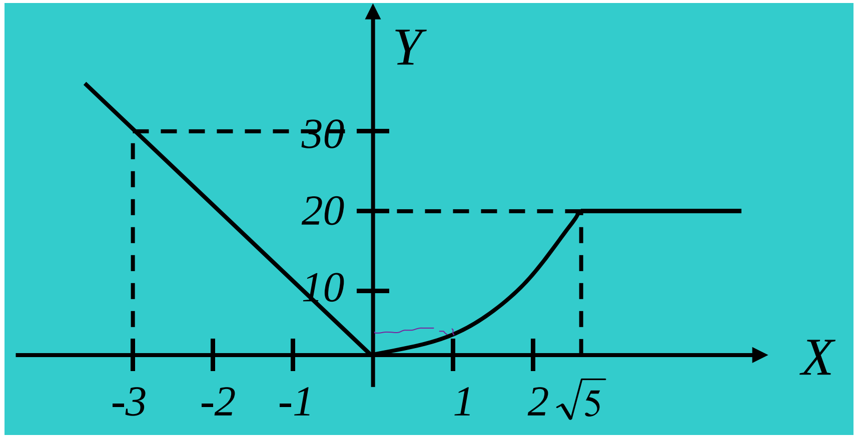
Same as previous $f_y(y)$ for $Y < 20$.

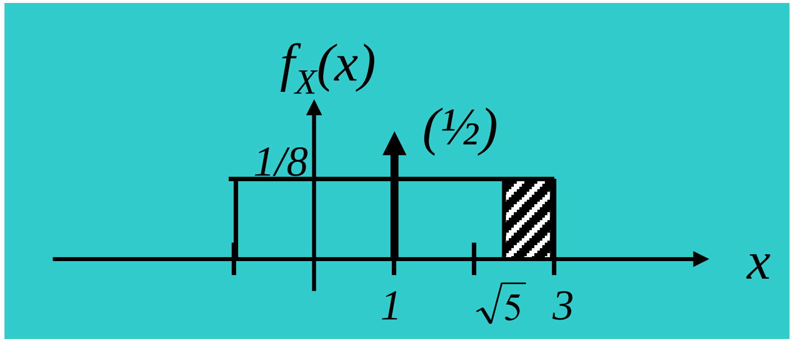
All X's from $\sqrt{5}$ to 3 gets mapped to $Y = 20$

$$P(Y = 20) = \frac{3-\sqrt{5}}{4}$$

Example – 3

Same as Ex 2 except $f_X(x)$ contains an impulse:





$P(Y=4) = \frac{1}{2}$.

$f_Y(y)$ same as Ex 2, except scaled by 1/2
AND the effect of impulse at $x = 1$

Short Summary

- Key points for the function of a RV
  - Find CDF $F_Y(y)$ and differentiate
  - Identify all values of X that map into that differential Y interval
  - Key equation $$f_Y(y) = \sum_{i=1}^{n} \frac{f_X(x_i)}{\left|\frac{dy}{dx}\right|_{x=x_i}}$$

- Special treatments for
  - Flat parts of $Y = G(X)$
  - Delta functions in $f_X(x)$

Expectation of a RV

Expectation of a Random Variable

Definition:

Discrete case: $$E(X) = \sum_i x_i p_X(x_i)$$

General case: $$E(X) = \int_{-\infty}^{+\infty} x f_X(x)\,dx$$

E(X) is well-defined if

$$\sum_i |x_i| p_X(x_i) < \infty$$

$$\int_{-\infty}^{+\infty} |x| f_X(x)\,dx < \infty$$

Interpretation and Notation

E(X) is a numerical average of a large number of independent observations of the random variable

E(X) is also known as the:

- first moment
- ensemble average
- mean

E(X) is symbolically expressed:

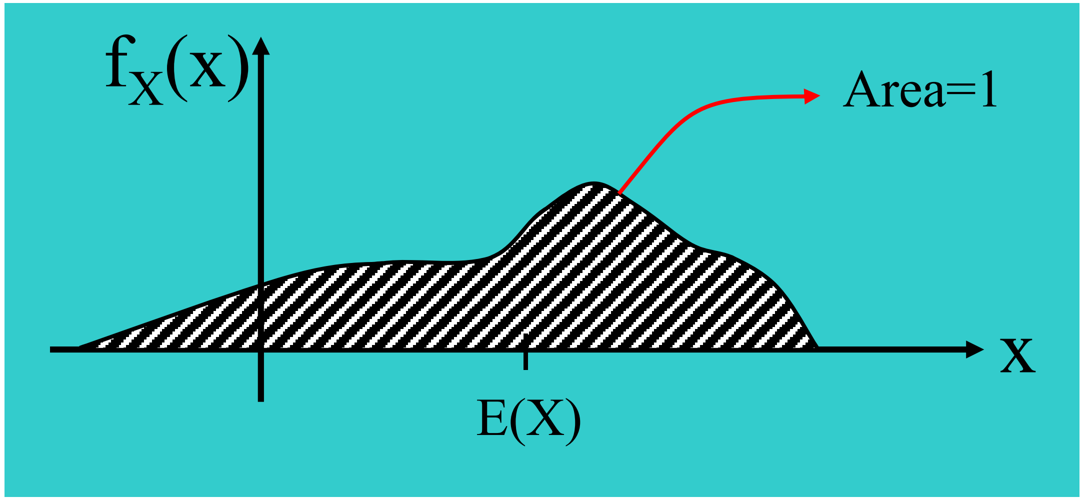
$$\mu_X, m_X, \eta_x, \text{or } \bar{X}$$

or just

$$\mu, m, \text{or } \eta$$

Interpretation

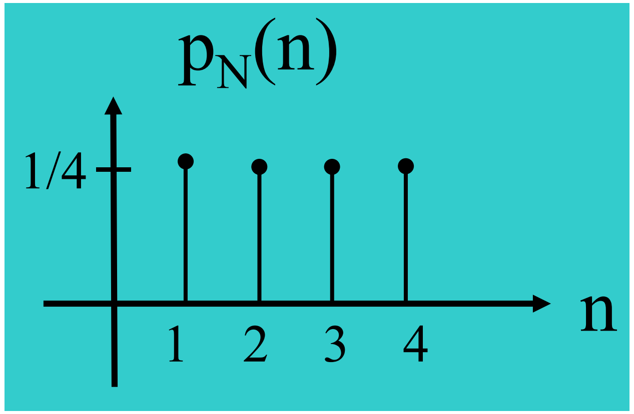
If the probability density is interpreted as a mass density along an axis, then $E(X)$ is the center of mass.



Note that $E(X)$ is not random.

Discrete Example

E(X) may not be a value that X can take.



$$E(N) = \sum_{n=1}^{4} n p_N(n) = 2.5$$

Mean of a Function of a RV

- To calculate $E\{G(X)\}$, there are two options:
  - First, get $f_Y(y)$ for $Y = G(X)$, then calculate $E(Y)$
  - Second, and faster, method: calculate

$$E[Y] = \sum_X G(x)p_X(x)$$

$$E[Y] = \int_{-\infty}^{+\infty} G(x)f_X(x)dx$$

- It is called the law of the unconscious statistician (LOTUS)

Proof of LOTUS

$$\boxed{\mathrm{P_r}(g(X)=y)=\sum_{x:g(x)=y}\mathrm{P_r}(X=x)=\sum_{x:g(x)=y}p_X(x)}$$

$$\mathrm{E}[Y]=\sum_y y p_Y(y)$$

$$=\sum_y y\,\mathrm{P_r}(Y=y)=\sum_y y\,\mathrm{P_r}(g(X)=y)$$

$$=\sum_y y\sum_{x:g(x)=y}p_X(x)$$

$$=\sum_x g(x)p_X(x)$$

Properties of Expected Value

1. The expected value of a constant* is that constant.

$$E(c) = c$$

2. The expected value is a linear operator:

$$E\big(cH(X)\big) = cE\big(H(X)\big), \quad c \in C$$

$$E\big(H(X) + G(X)\big) = E\big(H(X)\big) + E\big(G(X)\big)$$

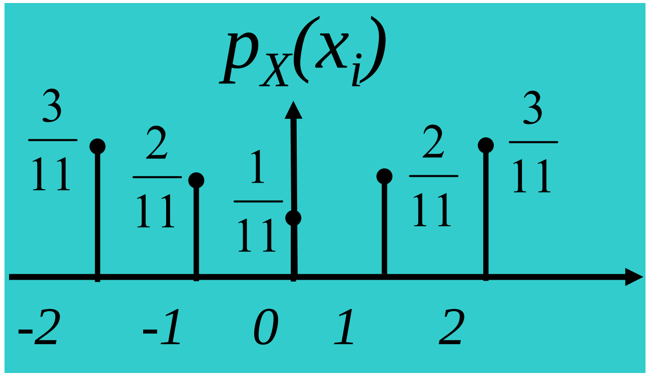
Ex:

$$Y = aX^2 + bX + c$$

$$\Rightarrow E(Y) = aE(X^2) + bE(X) + c$$

* Constant with respect to the random variables

Example Calculations of Expectation



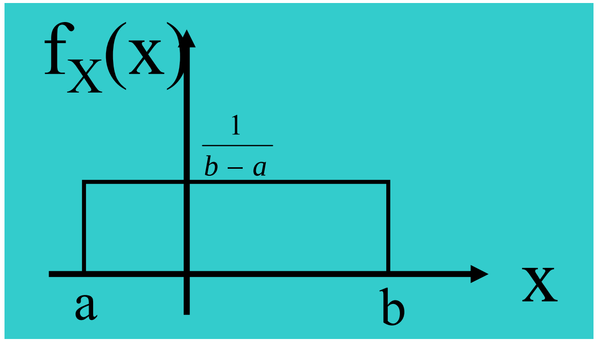
$$R_X : \{0, \pm 1, \pm 2\}$$

$$E(X^2)$$

Mean of a Uniform RV

E(X) is always in the middle of a uniform distribution.

$$E(X)=\int_{-\infty}^{+\infty} x f_X(x)\,dx=\int_a^b x\frac{1}{b-a}\,dx$$



$$=\frac{1}{b-a}\left(\frac{x^2}{2}\right)\Bigg|_a^b$$

$$=\frac{b^2-a^2}{2(b-a)}=\frac{b+a}{2}=\text{midpoint of ROS}$$

Expected Value of a Binomial RV

$$p_N(n) = \binom{m}{n} p^n (1-p)^{m-n}$$

Represent $N = \sum_{i=1}^{m} X_i$ $\quad X_i$ = Independent Bernoulli RV

$$E[N] = \quad m \cdot p$$

Expected Value of a Poisson RV

$$E(N) = \sum_{n=0}^{\infty} n \frac{\lambda^n}{n!} e^{-\lambda}$$

Change variables $i = n - 1$

$$E(N) = \lambda \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} e^{-\lambda} = \lambda$$

Mean of a Gaussian RV

$$E(X)=\int_{-\infty}^{+\infty} x\left(\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-m)^2}{2\sigma^2}}\right)dx$$

Let $y = x - m$. Then $x = y + m$ and $dx = dy$

$$E(X)=\int_{-\infty}^{+\infty}(y+m)\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{y^2}{2\sigma^2}}dy$$

$$=\int_{-\infty}^{+\infty}\underbrace{y\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{y^2}{2\sigma^2}}}_{\text{Odd}}dy+m\int_{-\infty}^{+\infty}\underbrace{\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{y^2}{2\sigma^2}}}_{\text{Just a PDF}}dy$$

Proof for the Mean

The mean is m, given that the first term is 0

$$\int_{-\infty}^{+\infty} y \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{y^2}{2\sigma^2}} dy$$

$$= \int_{-\infty}^{0} y \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{y^2}{2\sigma^2}} dy + \int_{0}^{+\infty} y \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{y^2}{2\sigma^2}} dy$$

Change of limits $$= -\int_{0}^{-\infty} y \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{y^2}{2\sigma^2}} dy + \int_{0}^{+\infty} y \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{y^2}{2\sigma^2}} dy$$

Change of variable $$= \int_{0}^{\infty} (-y) \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(-y)^2}{2\sigma^2}} dy + \int_{0}^{+\infty} y \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{y^2}{2\sigma^2}} dy$$

$$= 0$$

Variance

Observe that because E(X) is not random (a constant),

$$E\left[X+E(X)\right]=2E(X)$$

Suppose $\quad H(x)=(x-\mu_x)^2$

$=$ square of distance of X from it's mean

Definition for variance:

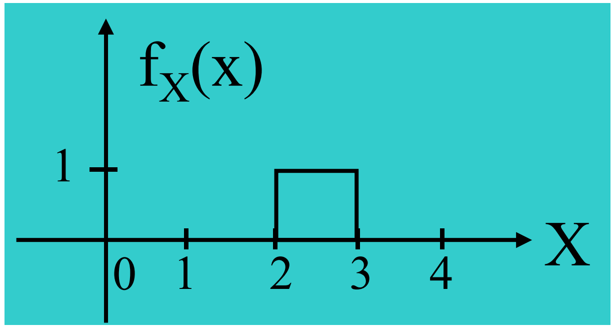
$$V(X)=E\left[H(X)\right]=E\left[(X-\mu_x)^2\right]$$

Alternative notation: $\mathrm{Var}(X)=\sigma_x^2$

Interpretation

- Observe that since $(X-\mu_x)^2$ is always positive, $V(X)$ must also be positive.
- The standard deviation, $\sqrt{\sigma_x^2}=\sigma_x$ is a measure of the width or spread of the PDF.

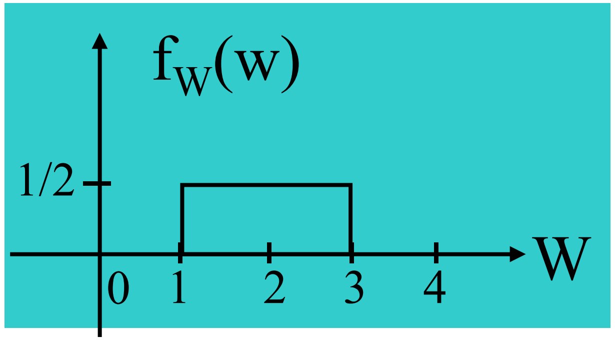
Example - I



$$\mu_X = 2.5$$

$$V(X)$$

Example - II



$$\mu_W = 2$$

$$V(W)$$

Alternative Formula

$$V(X)=E\left[(X-\mu_X)^2\right]=E\left(X^2-2X\mu_X+\mu_X{}^2\right)$$

$$=E\left(X^2\right)-2E(X)\mu_X+\mu_X{}^2=E\left(X^2\right)-\mu_X{}^2$$

or:

$$V(X)=E\left(X^2\right)-E(X)^2$$

Observe that if

$$\mu_x=0,\quad V(X)=E(X^2)$$

Variance of a Gaussian RV

Recall: $$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-m)^2}{2\sigma^2}}$$

$$V(X) = E\left[(X-m)^2\right] = \int_{-\infty}^{+\infty} \frac{(x-m)^2}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-m)^2}{2\sigma^2}}\, dx$$

$$= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} y^2 e^{-\frac{y^2}{2}}\, dy \qquad y = \frac{x-m}{\sigma}, \quad dy = \frac{dx}{\sigma}$$

Variance of a Gaussian RV, Concluded

Integration by parts:

$$u = y, \quad dv = ye^{-\frac{y^2}{2}}$$

$$du = dy, \quad v = -e^{-\frac{y^2}{2}}$$

$$V(X) = \frac{\sigma^2}{\sqrt{2\pi}}\left[-ye^{-y^2/2}\Big|_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} -e^{-y^2/2}dx\right]$$

$$= \frac{\sigma^2}{\sqrt{2\pi}}\left[0 + \sqrt{2\pi}\right] = \sigma^2$$

Almost a Gaussian PDF

Moments

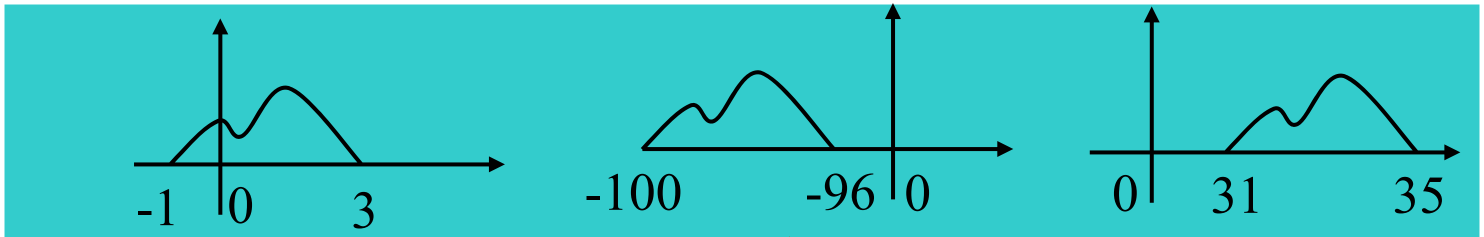
Definition: $\text{k}^{\text{th}}\text{ moment} = E(X^k)$

$$\text{k}^{\text{th}}\text{central moment} = E\left[(X-\mu_X)^k\right]$$

$$\text{k}^{\text{th}}\text{absolute moment} = E\left[|X|^k\right]$$

Observation:

These three PDFs have the same k^{th} central moment



Just shifted versions of the same function.

Short Summary

- Expectation of a RV $E(X) = \int_{-\infty}^{+\infty} x f_X(x)dx$
- Variance $V(X) = E\left[(X-\mu_x)^2\right]$ or $E\left(X^2\right) - E\left(X\right)^2$
- Moments
  - kth moment
  - kth central moment
  - kth absolute moment

Thank You!