Probability and Random Process

Aimin Tang

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

Sep. 27 2020

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.

$$X \sim N(0, \sigma^2)$$
$$Y \sim \text{Chi Square}$$

CDF for Function of a RV

- Find the corresponding set of X

  $\Pr(Y \in A) = \Pr(X \in G^{-1}(A)) = \Pr(X \in \{x : G(x) \in A\})$

- 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

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

$$F_Y(y) = \mathrm{P_r}(Y \le y) = \mathrm{P_r}\left(2e^{3X} \le y\right)$$

$$= \begin{cases} \mathrm{P_r}(X \le \frac{1}{3}\ln(\frac{y}{2})), & y > 0 \\ \mathrm{P}_X(\emptyset), & y \le 0 \end{cases}$$

consider all the possible values of $y \in \mathbb{R}$.

$$\boxed{= \begin{cases} F_X(\frac{1}{3}\ln(\frac{y}{2})), & y > 0 \\ 0, & y \le 0 \end{cases}}$$

PMF for discrete X

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
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.

We found cdf of Y previously

$$F_Y(y) = \begin{cases} F_X(\frac{1}{3}\ln(\frac{y}{2})), & y > 0 \\ 0, & y \leq 0 \end{cases}$$

Example-cont

Plug in $f_X(x)$

$$\frac{1}{3}\ln(\frac{y}{2}) \in [-1, 1] \Rightarrow y \in [2e^{-3}, 2e^{3}]$$

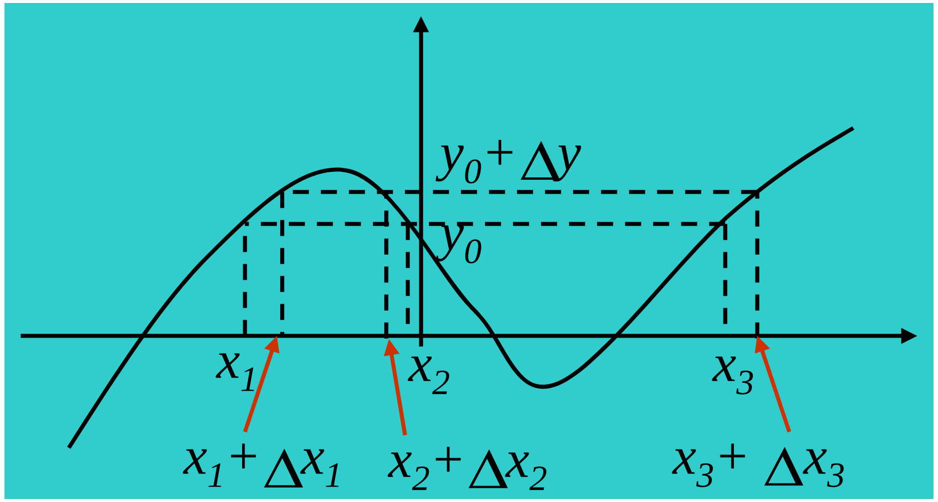
$$f_Y(y) = \begin{cases} \frac{1}{6y}, & y \in [2e^{-3}, 2e^{3}] \\ 0, & \text{otherwise} \end{cases}$$

The method of differentials - I

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

$$y_0 \leq Y \leq y_0 + \Delta y$$

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



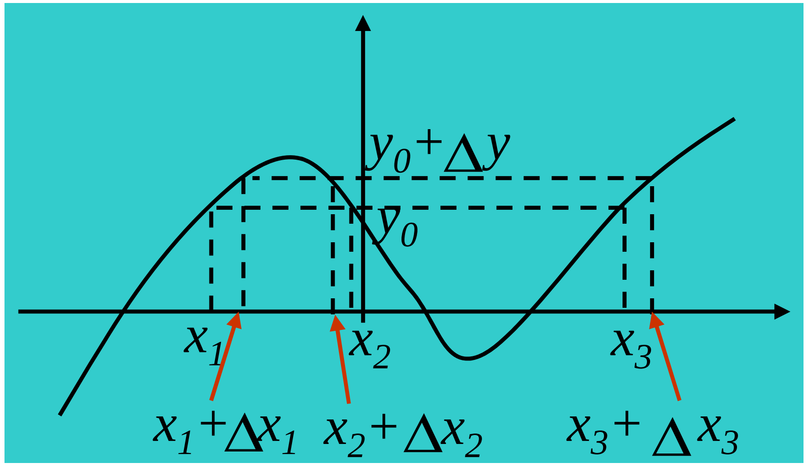
x_1, x_2 , and x_3 are solutions to

$Y=G(X)$

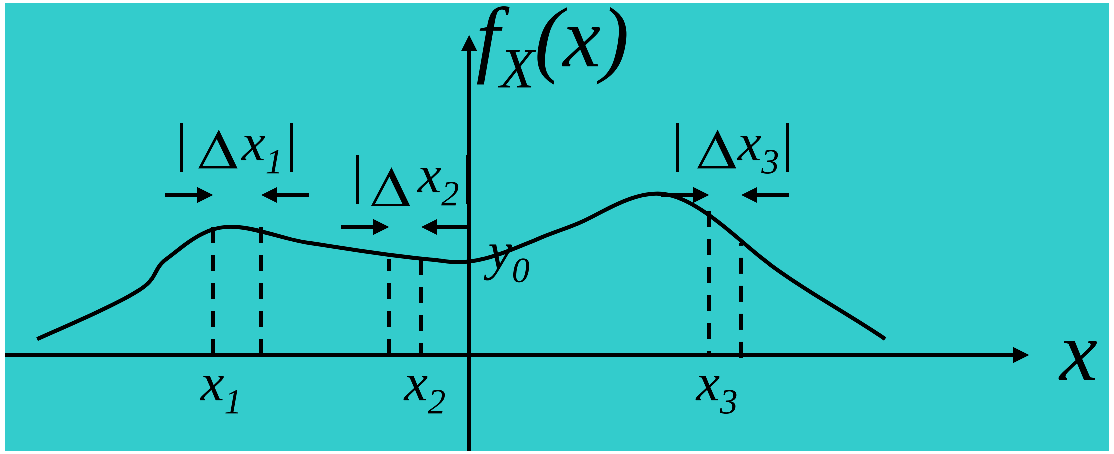
The method of differentials - II

$$
\begin{aligned}
P(y_0 \le Y \le y_0 + \Delta y) = P(&x_1 \le X \le x_1 + \Delta x_1 \\
&\cup x_2 \le X \le x_2 + \Delta x_2 \\
&\cup x_3 \le X \le x_3 + \Delta x_3)
\end{aligned}
$$

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



The method of differentials - III



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

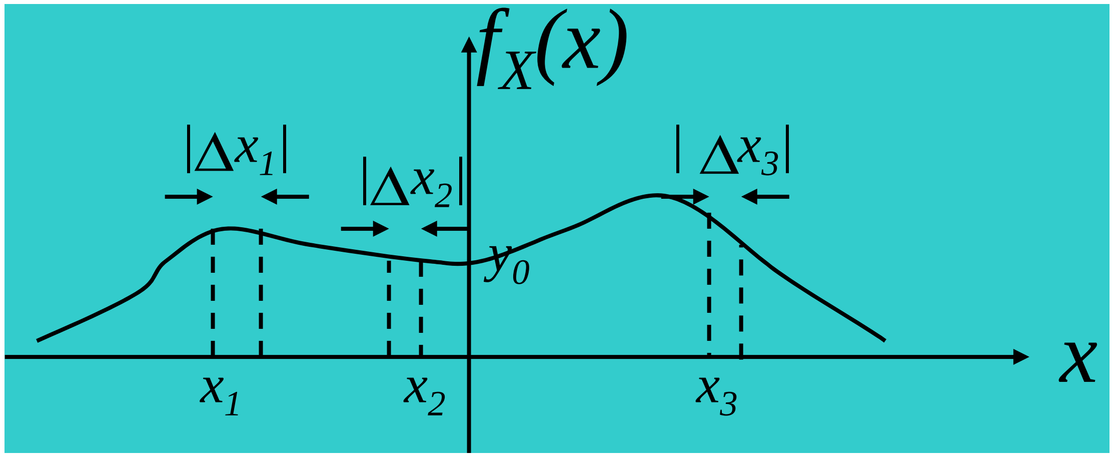
$$P(x_i \le X \le x_i + \Delta x_i) \approx f(x_i)\Delta x_i$$

The method of differentials - IV

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

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

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



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,

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

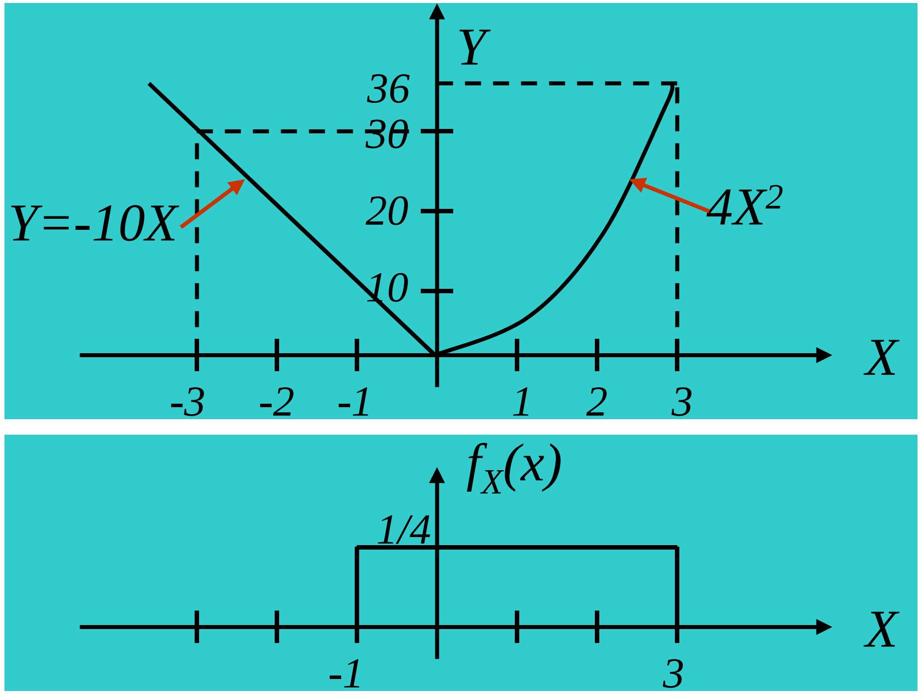
Where n is the number of solutions to $Y = G(X)$.

★REMEMBER DO NOT APPLY TO

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

Function of a RV Examples

Ex. 1:



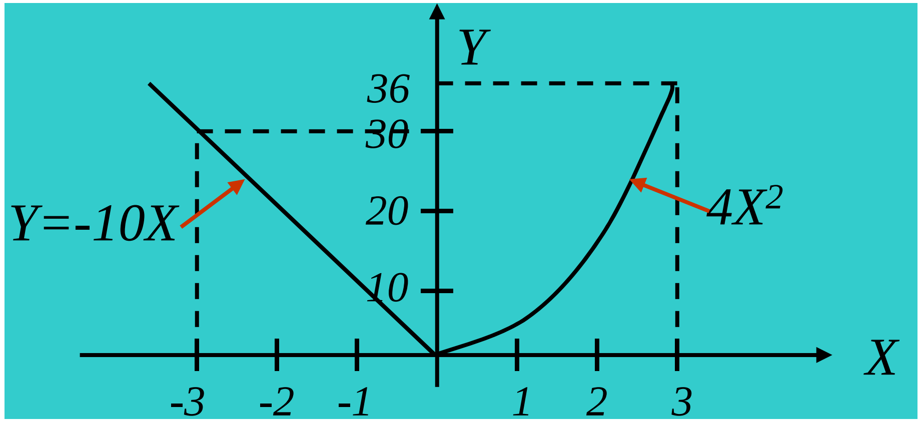
Observe that $f_Y(y) = 0$ for $y > 36$ and $y < 0$ because no probability mass maps to these regions.

Example, Continued

For $y > 0$, there are two solutions:

$$x_1 = -\frac{y}{10}$$

$$x_2 = \frac{\sqrt{y}}{2}$$

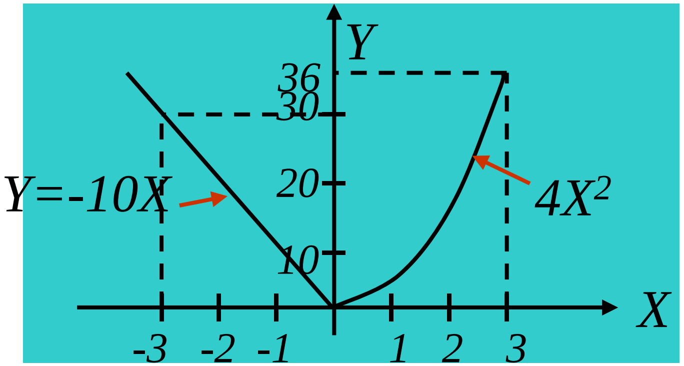


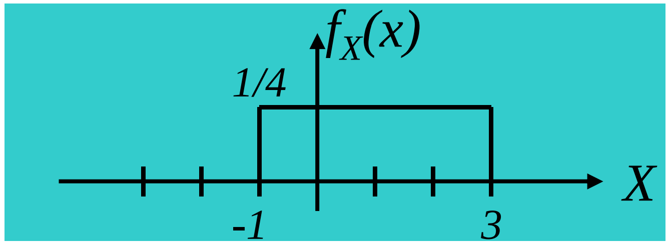
The slopes for these solutions are

$$\text{at } x_1 : \frac{dy}{dx} = -10 \qquad \text{at } x_2 : \frac{dy}{dx} = 8x$$

Example, Continued

Since $f_X(x) = 0$ for $x_1 < -1, x_1$ contributes to the answer only when $x_1 > -1$ or when $y \le 10$



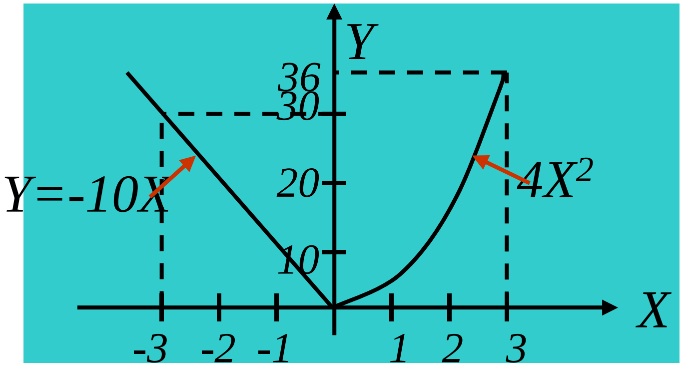


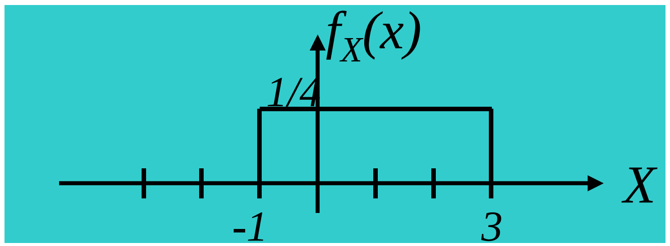
$$
f_Y(y) = \begin{cases} 0 & y > 36 \text{ and } y \le 0 \\ \dfrac{f_X\left(\dfrac{\sqrt{y}}{2}\right)}{\left|8\dfrac{\sqrt{y}}{2}\right|} & 10 < y \le 36 \\ \dfrac{f_X\left(\dfrac{\sqrt{y}}{2}\right)}{\left|8\dfrac{\sqrt{y}}{2}\right|} + \dfrac{f_X\left(\dfrac{-y}{10}\right)}{|-10|} & 0 < y \le 10 \end{cases}
$$

Example, Continued

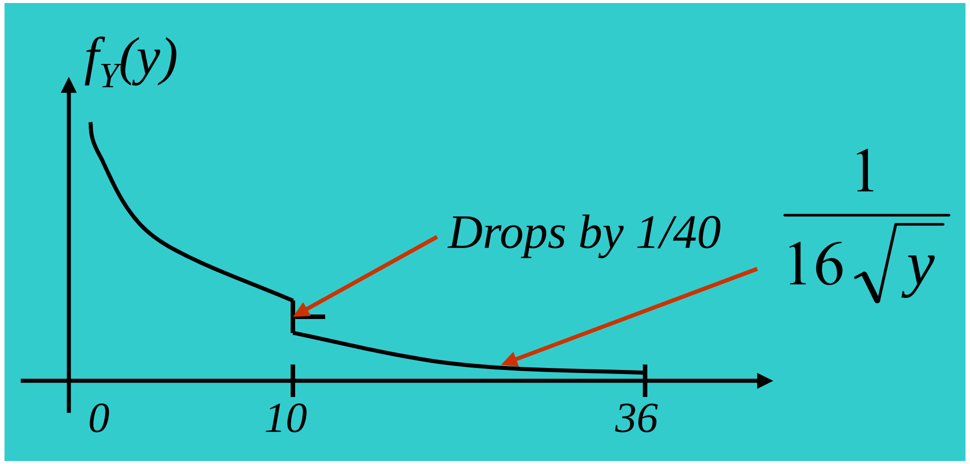
Plug in $f_X(x)$ function:

$$f_Y(y) = \begin{cases} 0 & y > 36 \text{ and } y \le 0 \\ \dfrac{1}{16\sqrt{y}} & 10 < y \le 36 \\ \dfrac{1}{16\sqrt{y}} + \dfrac{1}{40} & 0 < y \le 10 \end{cases}$$



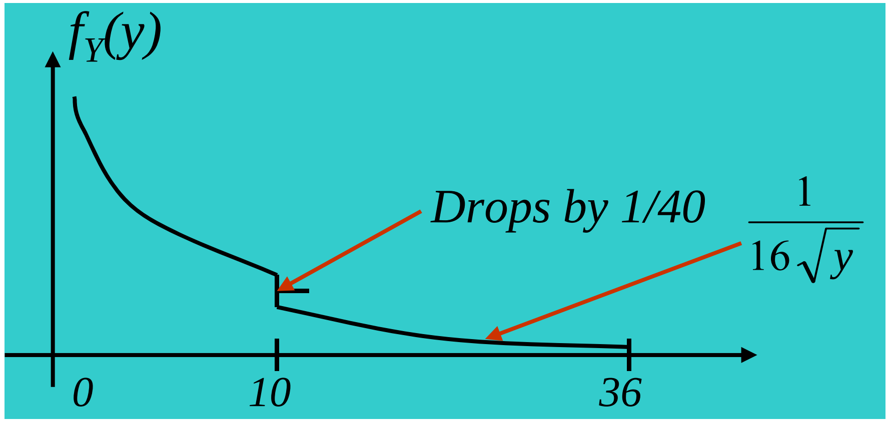


Example, Concluded



Check that $\int_{-\infty}^{+\infty} f_Y(y)dy = 1$

Check

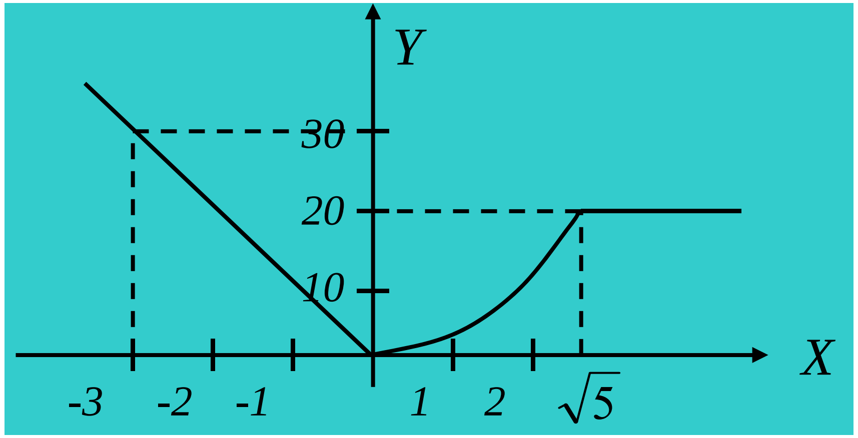


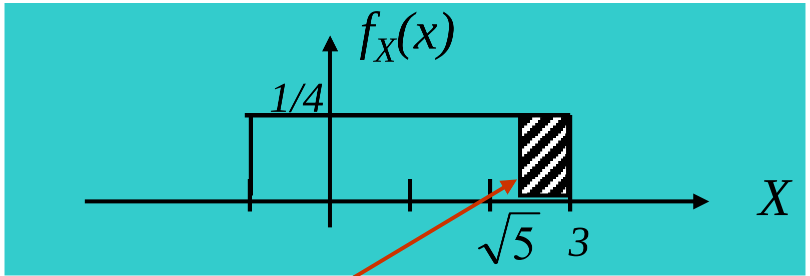
$$\int_0^{10}\left(\frac{1}{16\sqrt{y}}+\frac{1}{40}\right)dy+\int_{10}^{36}\frac{1}{16\sqrt{y}}dy$$

$$=\left(\frac{\sqrt{y}}{8}+\frac{y}{40}\right)\Bigg|_0^{10}+\left(\frac{\sqrt{y}}{8}\right)\Bigg|_{10}^{36}=\frac{\sqrt{10}}{8}+\frac{1}{4}+\frac{\sqrt{36}}{8}-\frac{\sqrt{10}}{8}=1$$

Example – 2

Same as Ex 1 but function has a flat part:





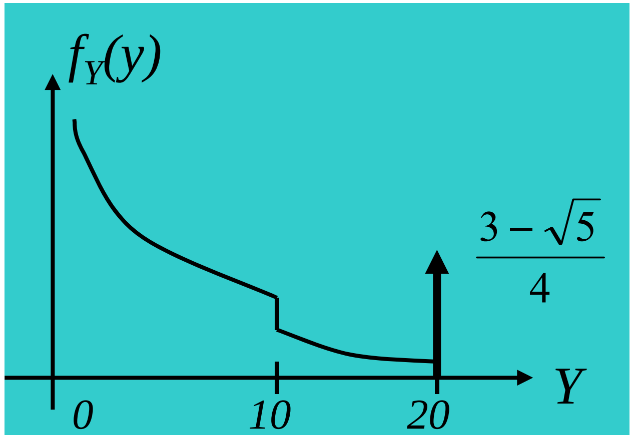
Shaded area= $\dfrac{3-\sqrt{5}}{4}$

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

All X's from √5 to 3 gets mapped to $Y = 20$

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

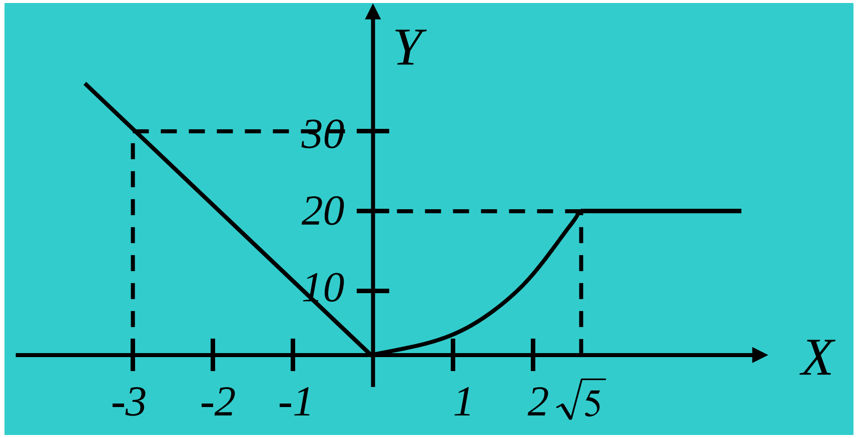
Example – 2, Concluded

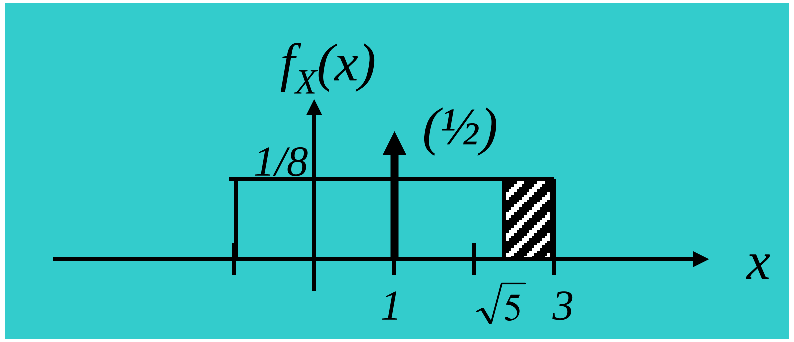


$$f_Y(y) = \begin{cases} 0 & y < 0 \text{ and } y > 20 \\ \dfrac{1}{16\sqrt{y}} + \dfrac{1}{40} & 0 \le y \le 10 \\ \dfrac{1}{16\sqrt{y}} + \left(\dfrac{3-\sqrt{5}}{4}\right)\delta(y-20) & 10 < y \le 20 \end{cases}$$

Example – 3

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

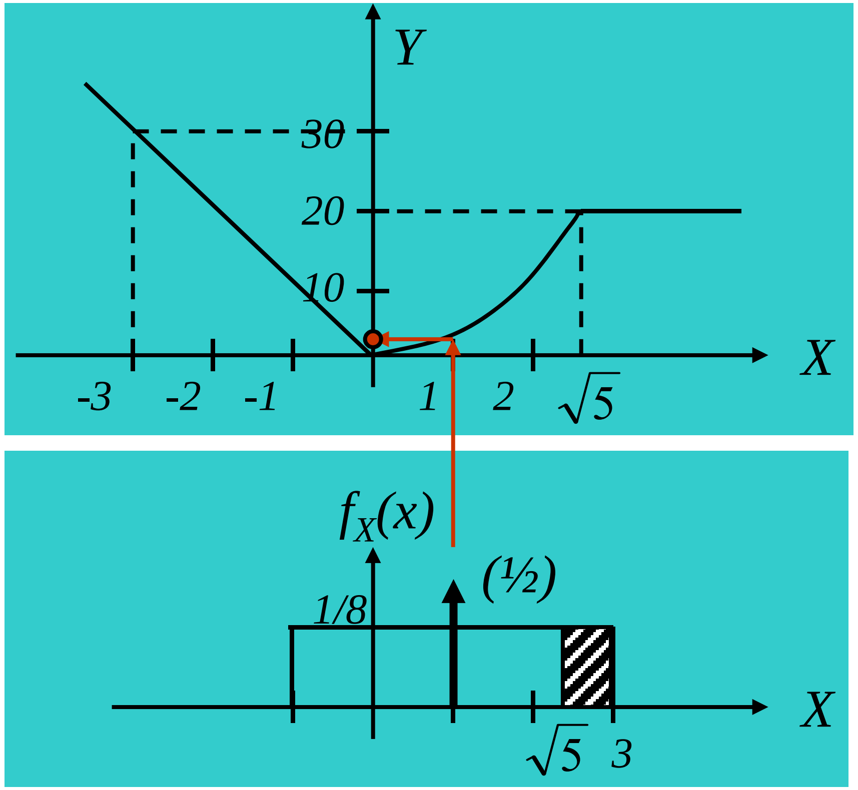




$f_X(x)$ same as Ex 2, except scaled by 1/2

AND the effect of impulse at $x = 1$

Example – 3, Concluded

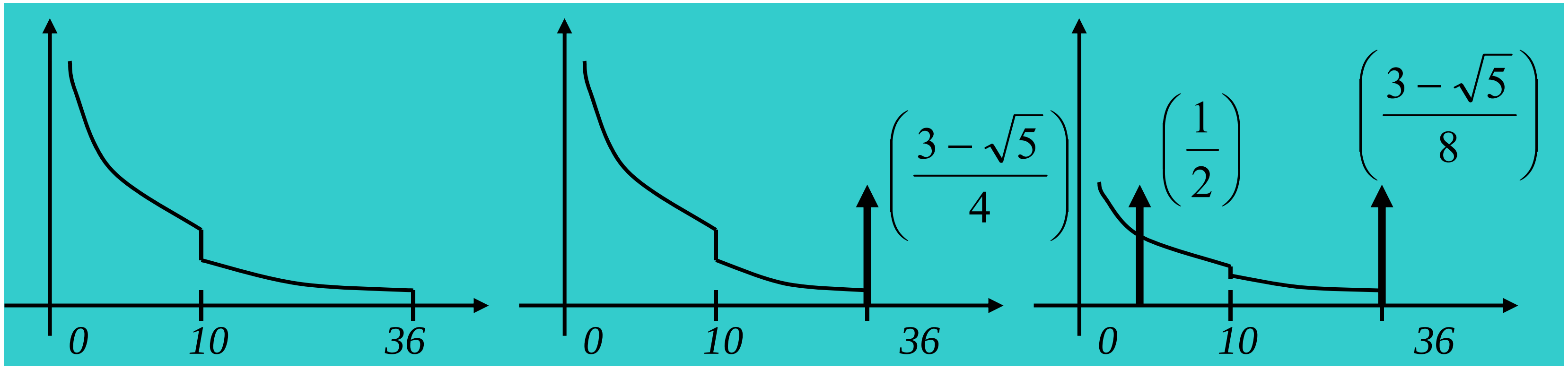


$f_X(x)$ same as Ex 2, except scaled by $1/2$

The prob. of $1/2$ at $x = 1$ is mapped directly to $y = 4(1)^2$, yielding an impulse in $f_Y(y)$ of prob. $1/2$ at $y = 4$.

Comparison of $f_Y(y)$

Ex 1: Ex 2: Ex 3:



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:

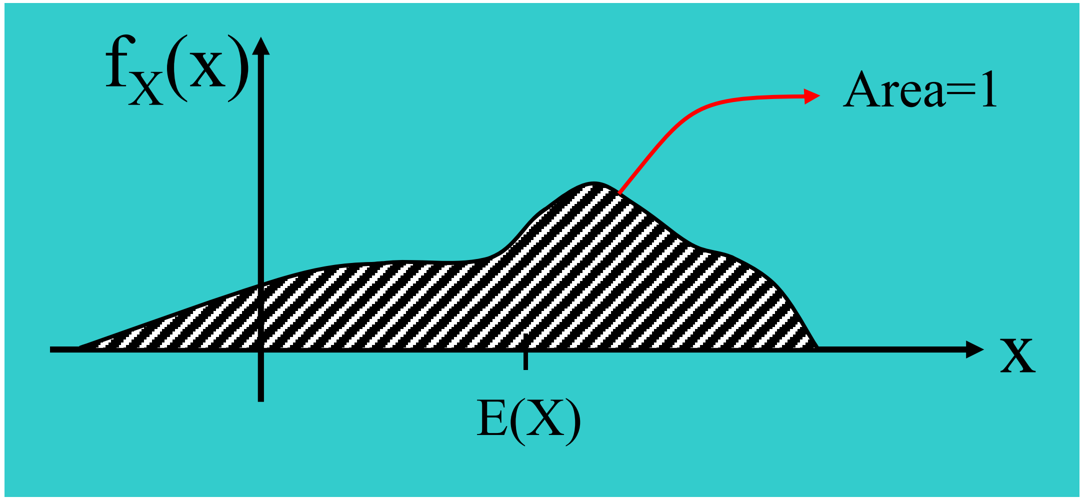
$$\mu_X, m_X, \eta_x, \text{or } \overline{X}$$

or just

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

Interpretation

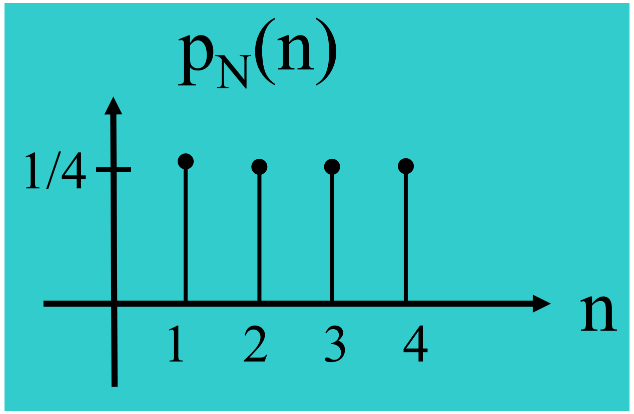
If the probability density is interpreted as a mass density along an axis, then $E(X)$ is the center of mass.



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

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

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

$$= \sum_y y \Pr(Y = y) = \sum_y y \Pr(g(X) = y)$$

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

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

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:

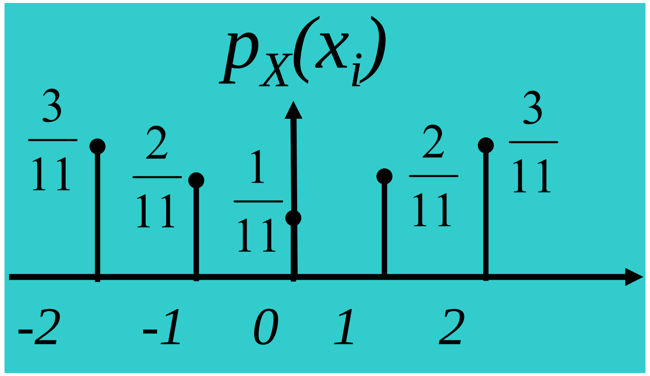
$$E\big(cH(X)\big) = cE\big(H(X)\big), \quad c \in C$$
$$E\big(H(X) + G(X)\big) = E\big(H(X)\big) + E\big(G(X)\big)$$

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



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

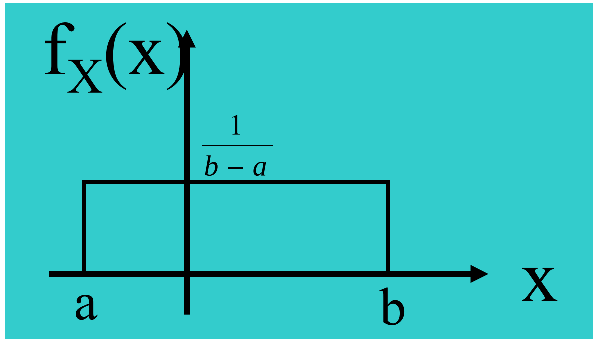
$$E(X^2) = \sum_{i=-2}^{2} i^2 p_X(x_i)$$ by LOTUS

$$= 0 \cdot \frac{1}{11} + 2\left(1^2 \cdot \frac{2}{11} + 2^2 \cdot \frac{3}{11}\right) = \frac{28}{11} = 2.54$$

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

$$E[N] = E\left[\sum_{i=1}^{m} X_i\right] = \sum_{i=1}^{m} E(X_i) = \sum_{i=1}^{m} p = mp$$

Mean of a sum is the sum of the means

Expected Value of a Poisson RV

$$E(N)=\sum_{n=0}^{\infty} n\frac{\lambda^n}{n!}e^{-\lambda}=\sum_{n=1}^{\infty} n\frac{\lambda^n}{n!}e^{-\lambda}=\lambda\sum_{n=1}^{\infty}\frac{\lambda^{n-1}}{(n-1)!}e^{-\lambda}$$

Dropped $n = 0$

Change variables $i = n - 1$

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

poisson
PMF

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 $dx = dy$

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

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

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

$y' = -y.$

$$= 0$$

Variance

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

$$E[X + E(X)] = 2E(X)$$

μ

$E[X] + \mu = 2E[X]$

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

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

Definition for variance:

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

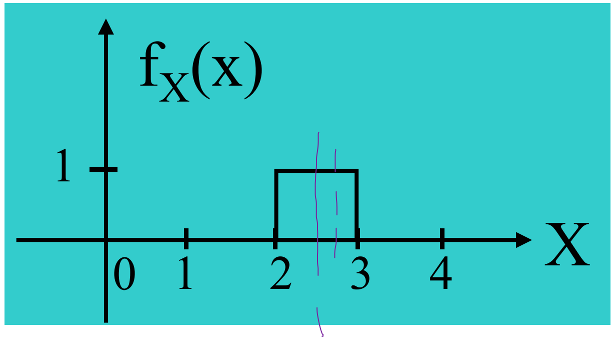
$\mathrm{Var}[X]$

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



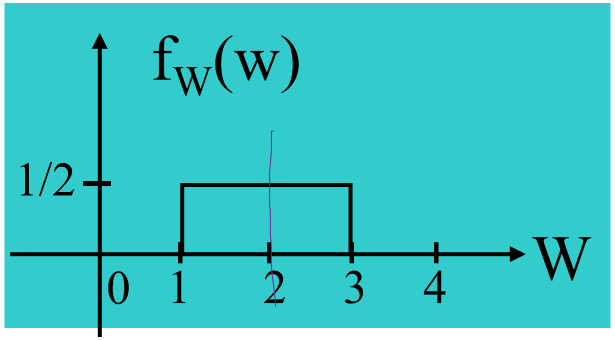
$$\mu_X = 2.5$$

$$V(X) = \int_2^3 (x-2.5)^2 \cdot 1dx = \int_2^3 \left(x^2 - 5x + (2.5)^2\right)dx$$

$$= \left(\frac{x^3}{3} - \frac{5x^2}{2} + (2.5)^2 x\right)\Bigg|_2^3 = \frac{27-8}{3} - \frac{5(9-4)}{2} + (2.5)^2(3-2)$$

$$= \frac{19}{3} - \frac{25}{2} + \frac{25}{4} = \frac{76-150+75}{12} = \frac{1}{12}$$

Example - II



range' → 2 range.

variance' → 4 variance.

$$\mu_W = 2$$

$\int (w-\mu)^2 f(w)\, dw.$

$w^2 - 4w + 4.$

$$V(W) = \int_1^3 (w-2)^2 \cdot \frac{1}{2} dw = \frac{1}{2}\left(\frac{w^3}{3} - 2w^2 + 4w\right)\Bigg|_1^3$$

$$= \frac{1}{2}\left[\frac{27-1}{3} - 2(9-1) + 4 + (3-1)\right] = \frac{1}{2}\left[\frac{26}{3} - 16 + 8\right] = \frac{1}{3}$$

$\int_1^3 w^2 \cdot \frac{1}{2} dw = \frac{1}{2} \cdot \frac{1}{3} w^3 \Big|_1^3 = \frac{1}{6}(27-1) = \frac{13}{3}$

$- \left(\int_1^3 w \cdot \frac{1}{2} dw\right)^2 = \frac{1}{2} \cdot \frac{1}{2} w^2 \Big|_1^3 = \left(\frac{1}{4}(8)\right)^2 = 4$

$\frac{13}{3} - 4 = \frac{1}{3}$

Alternative Formula

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

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

or:

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

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:

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

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

$$V(X) = \frac{\sigma^2}{\sqrt{2\pi}}\left[-ye^{-y^2/2}\Big|_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dx \right]$$

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

Almost a Gaussian PDF

Moments

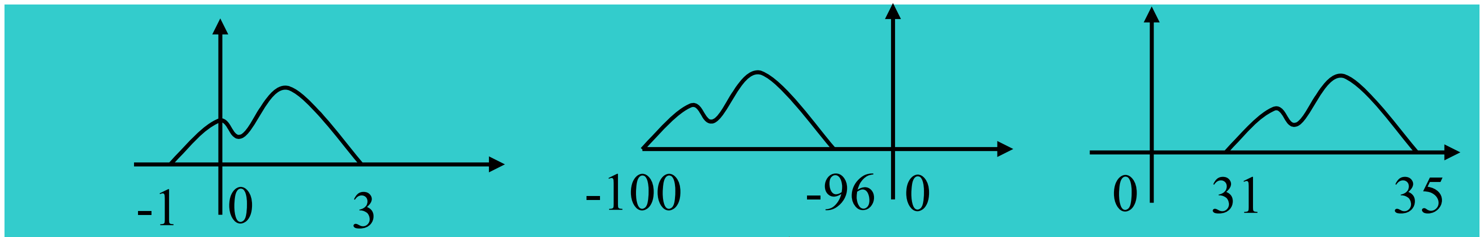
Definition: $\text{k}^{\text{th}}\ \text{moment} = E(X^k)$

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

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

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!