Probability and Random Processes ECS 315

Asst. Prof. Dr. Prapun Suksompong prapun@siit.tu.ac.th 10 Continuous Random Variables



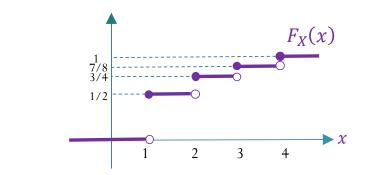
Office Hours:

Check Google Calendar on the course website. Dr.Prapun's Office: 6th floor of Sirindhralai building, BKD

Sections 10.1-10.2

Discrete RV

- **pmf**: $\boldsymbol{p}_X(x) \equiv P[X = x]$
 - Two characterizing properties:
 - $p_X(x) \ge 0$
 - $\sum_{x} p_X(x) = 1$
- $S_X = \{x: p_X(x) > 0\}$
- *P*[some statement(s) about *X*]
 - $= \sum_{\substack{\{\text{all the } x \text{ values that} \\ \text{satisfy the statement(s)}\}}} p_X(x)$
- **cdf** is a staircase function with jumps whose size at x = c gives P[X = c].



Continuous RV

• P[X = x] = 0

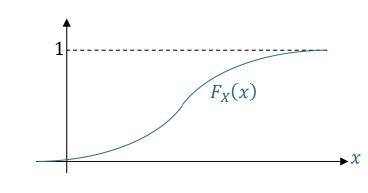
probability per unit length

- $\mathbf{pdf}: P[x_0 \le x \le x_0 + \Delta x] \approx \mathbf{f}_X(x_0) \Delta x$
 - Two characterizing properties:
 - $f_X(x) \ge 0$
 - $\int_{-\infty}^{\infty} f_X(x) \, dx = 1$
- $S_X = \{x: f_X(x) > 0\}$
- *P*[some statement(s) about *X*] =

$$f_X(x)dx$$

{all the x values that satisfy the statement(s)}

• **cdf** is a continuous function.



Chapter 9 vs. Section 10.3Discrete RVContinuous RV
$$\mathbb{E}X = \sum_{x} xp_{x}(x)$$
 $\mathbb{E}X = \int_{-\infty}^{\infty} xf_{x}(x)dx$ $\mathbb{E}[g(X)] = \sum_{x} g(x)p_{x}(x)$ $\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x)f_{x}(x)dx$ $\mathbb{E}[X^{2}] = \sum_{x} x^{2}p_{x}(x)$ $\mathbb{E}[x^{2}] = \int_{-\infty}^{\infty} x^{2}f_{x}(x)dx$ $\operatorname{Var}[X] = \mathbb{E}[(X - \mathbb{E}X)^{2}] = \mathbb{E}[X^{2}] - (\mathbb{E}X)^{2}$ $\sigma_{x} = \sqrt{\operatorname{Var}[X]}$

Probability and Random Processes ECS 315

Asst. Prof. Dr. Prapun Suksompong prapun@siit.tu.ac.th 10.1 Probability Density Function

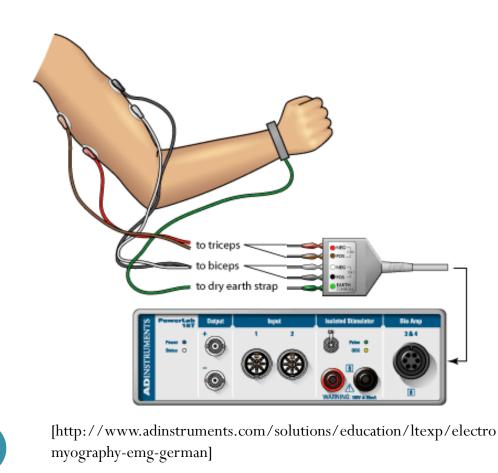
Ex. rand function

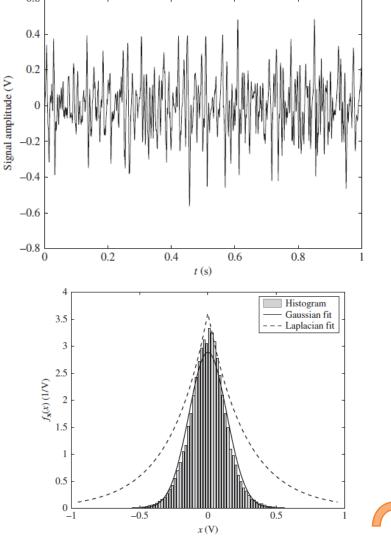
- Generate an array of uniformly distributed pseudorandom numbers.
 - The pseudorandom values are drawn from the standard uniform distribution on the open interval (0,1).
- rand returns a scalar.
- rand(m,n) or rand([m,n]) returns an *m*-by-*n* matrix.
 - rand(n) returns an *n*-by-*n* matrix

>> rand				
ans =				
0.3816				
>> rand(10,2)				
ans =				
0.7655	0.6551			
0.7952	0.1626			
0.1869	0.1190			
0.4898	0.4984			
0.4456	0.9597			
0.6463	0.3404			
0.7094	0.5853			
0.7547	0.2238			
0.2760	0.7513			
0.6797	0.2551			

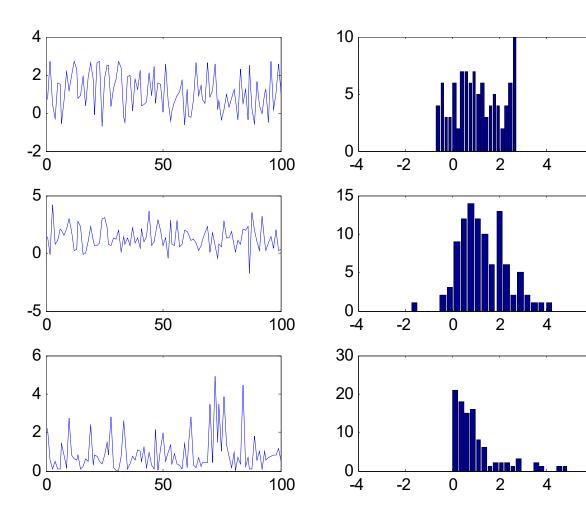
Ex. Muscle Activity

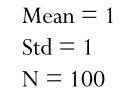
• Look at electrical activity of skeletal muscle by recording a human electromyogram (EMG).

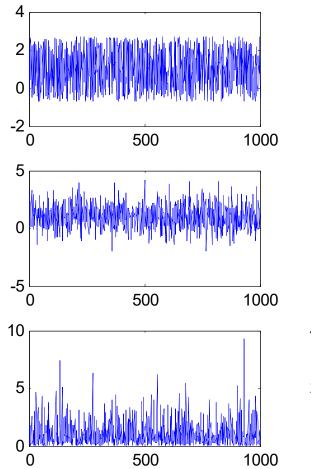


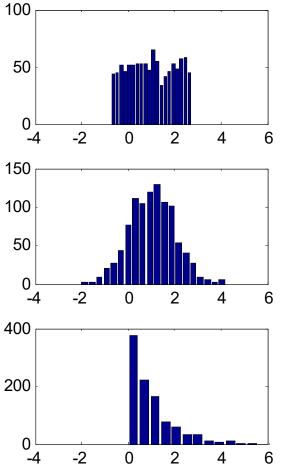


```
close all; clear all;
N = 1e6; b = 20; m = 1; s = 1;
R = [1-5*s, 1+5*s];
% Uniform
X = (2*sqrt(3)*(rand(1,N)-0.5))+1;
subplot(3,2,1); plot(X);
subplot(3,2,2); plotHistPdf(X,b)
xlim(R)
% Normal
X = randn(1, N) + 1;
subplot(3,2,3); plot(X);
subplot(3,2,4); plotHistPdf(X,b)
xlim(R)
% Exponential
X = exprnd(1, 1, N);
subplot(3,2,5); plot(X);
subplot(3,2,6); plotHistPdf(X,b)
xlim(R)
```

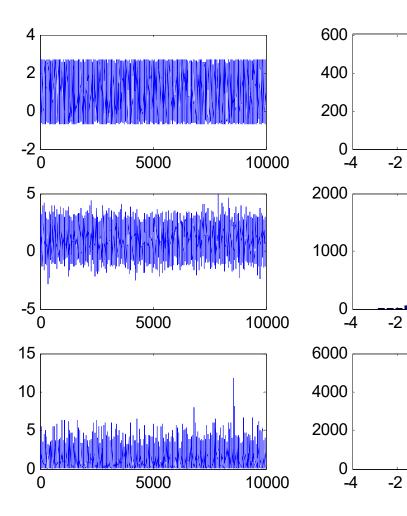


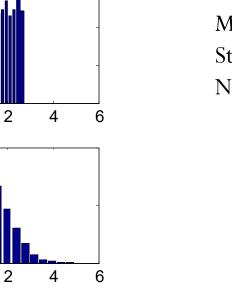






Mean = 1Std = 1N = 1,000





Mean = 1 Std = 1 N = 10,000

Review: *P*[some condition(s) on *X*]

For discrete random variable,

8.14. Steps to find probability of the form P [some condition(s) on X] when the pmf $p_X(x)$ is known.

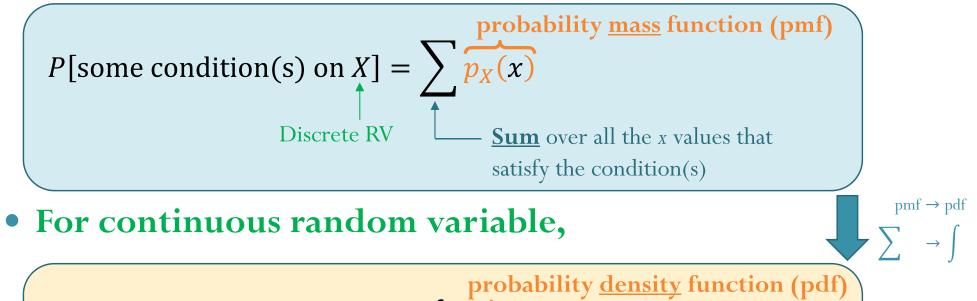
- (a) Find the support of X.
- (b) Consider only the x inside the support. Find all values of x that satisfy the condition(s).
- (c) Evaluate the pmf at x found in the previous step.
- (d) Add the pmf values from the previous step.

$$P[\text{some condition(s) on } X] = \sum_{x} p_X(x)$$

Discrete RV Sum over all the *x* values that satisfy the condition(s)

P[some condition(s) on *X*]

• For discrete random variable,

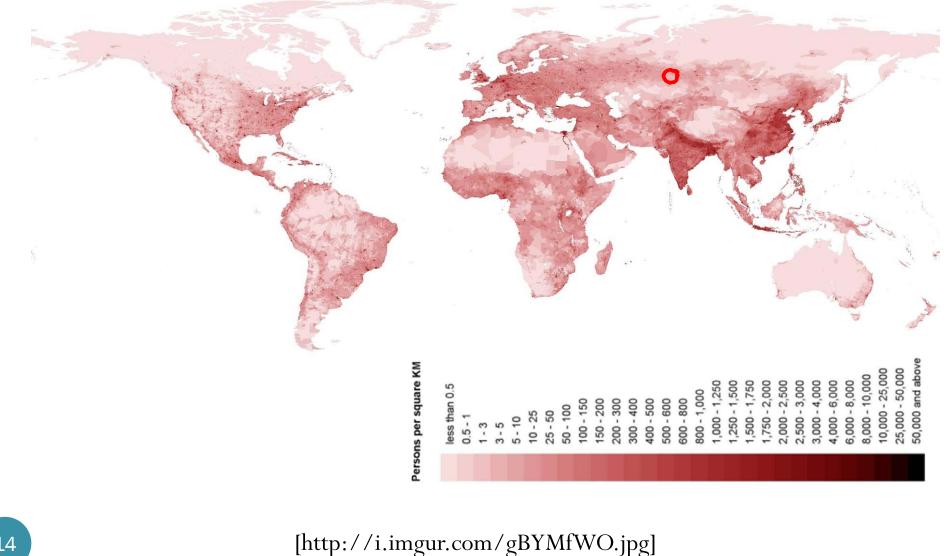


$$P[\text{some condition(s) on } X] = \int_{RV} \int_{RV} \int_{RV} \int_{RV} \int_{RV} \int_{RV} \int_{V} \int_$$

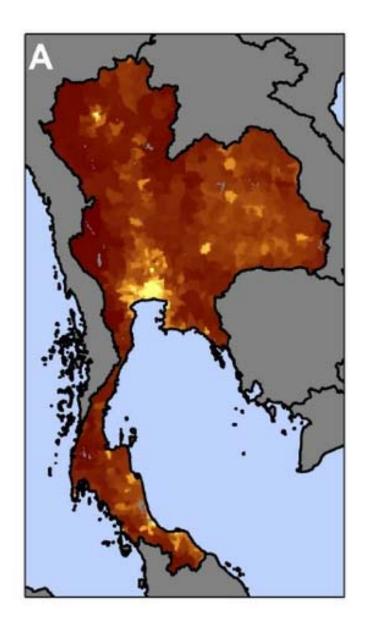
Support of a RV

- In general, the **support** of a RV X is any set S such that $P[X \in S] = 1$.
- In this class, we try to find the smallest (minimal) set that works as a support.
- For discrete random variable, $S_X = \{x: p_X(x) > 0\}$
- For continuous random variable, $S_X = \{x: f_X(x) > 0\}$

World Map of Population Density



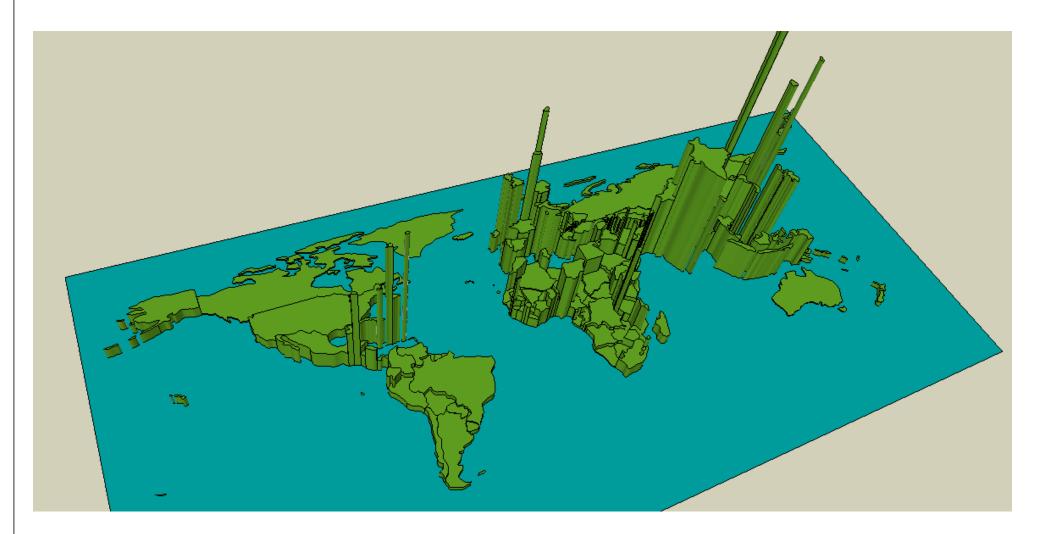
Thailand's Population Density



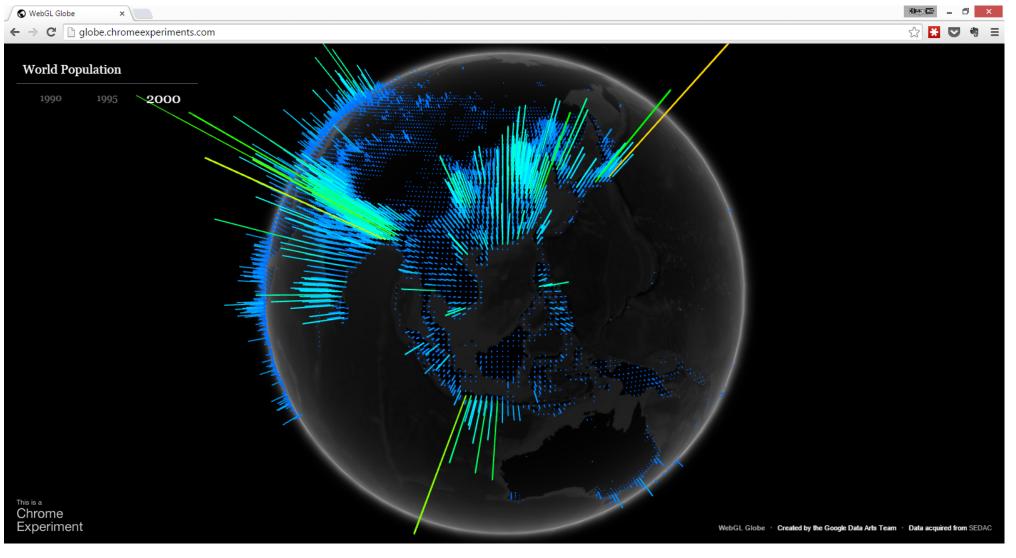


https://www.researchgate.net/pu blication/260378246 Climate-Related Hazards A Method for Global Assessment of Urban an d Rural Population Exposure to Cyclones Droughts and Floods /figures?lo=1

World Map of Population Density



World Map of Population Density



http://globe.chromeexperiments.com/

"Density"

- Density = quantity per unit of measure.
- Population Density = number of people per unit area
 - Location with high density value means there are a lot of people around that location.
 - Given a region, we integrate the density over that region to get the number of people residing in that region.
- Probability Density = probability per unit "length".
 - Given an interval, we integrate the density over that interval to get the probability that the RV will be in that interval.

References

- From Discrete to Continuous Random Variables: [Y&G]
 Sections 3.0 to 3.1
- PDF and CDF: [Y&G] Sections 3.1 to 3.2
- Expectation and Variance: [Y&G] Section 3.3
- Families of Continuous Random Variables: [Y&G] Sections 3.4 to 3.5

Course Outline

The following is a tentative list of topics with their corresponding chapters from the text: Yates and Goodman. Each topic spans approximately one week.

1.	Introduction, Set Theory, Classical Probability	[1]
2.	Combinatorics: Four Principles and Four Kinds of Counting Problems	[1]
3.	Probability Foundations	[1]
4.	Event-based Conditional Probability	[1]
5.	Event-based Independence	[1]
6.	Random variables, Support, Probability Distribution	[2]
7.	MIDTERM: 3 Oct 2019 TIME 15:00 - 17:00	
8.	Discrete Random Variables	[2]
9.	Families of Discrete Random Variables and Introduction to Poisson Processes	[2,10]
10.	Real-Valued Functions of a Random Variable	[2]
11.	Expectation, Moment, Variance, Standard Deviation	[2]
12.	Continuous Random Variables	[3]
13.	Families of Continuous Random Variables and Introduction to	[3,10]
	Poisson Processes	

• Exercise 17 Solution [Posted @ 5PM on C

- References: [Y&G] Chapter 2
- Notes from the tutorial session [Posted @ 11:3
- Part IV: Continuous Random Variables
 - Chapter 10: Continuous Random Variables [Pos
 - References
 - From Discrete to Continuous Rand
 - PDF and CDF: [Y&G] Sections 3.1 to
 - Expectation and Variance: [Y&G] Se
 - Families of Continuous Random Va

Probability and Random Processes ECS 315

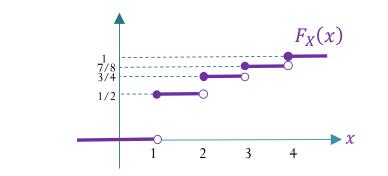
Asst. Prof. Dr. Prapun Suksompong prapun@siit.tu.ac.th 10.2 Properties of PDF and CDF

Sections 10.1-10.2

Discrete RV

21

- **pmf**: $\boldsymbol{p}_X(x) \equiv P[X = x]$
 - Two characterizing properties:
 - $p_X(x) \ge 0$
 - $\sum_{x} p_X(x) = 1$
- $S_X = \{x: p_X(x) > 0\}$
- *P*[some condition(s) on *X*]
 - $= \sum_{\substack{\{\text{all the } x \text{ values that} \\ \text{satisfy the condition(s)}\}}} p_X(x)$
- **cdf** is a staircase function with jumps whose size at x = c gives P[X = c].



Continuous RV

• P[X = x] = 0

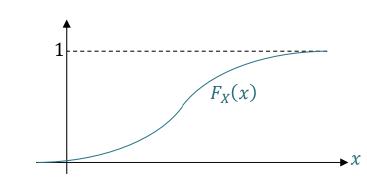
probability per unit length

- $\mathbf{pdf}: P[x_0 \le x \le x_0 + \Delta x] \approx \mathbf{f}_X(x_0) \Delta x$
 - Two characterizing properties:
 - $f_X(x) \ge 0$
 - $\int_{-\infty}^{\infty} f_X(x) \, dx = 1$
- $S_X = \{x: f_X(x) > 0\}$
- *P*[some condition(s) on *X*] =

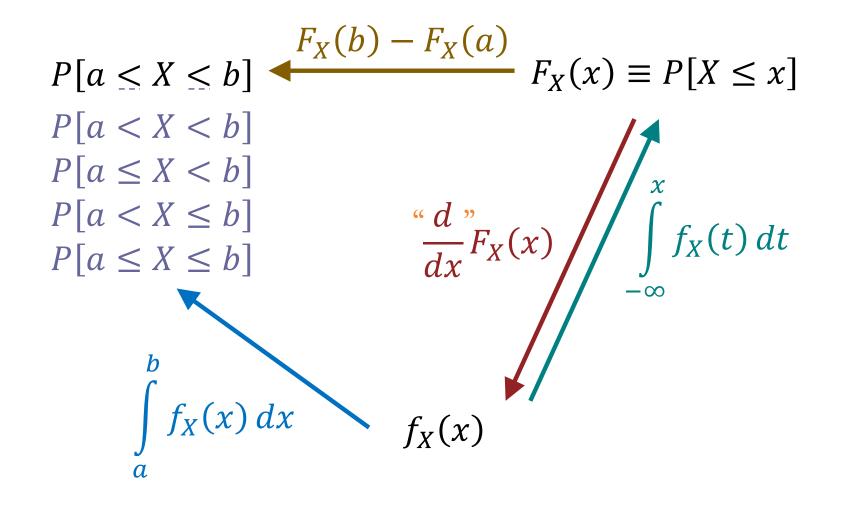
$$f_X(x)dx$$

{all the *x* values that satisfy the condition(s)}

• **cdf** is a continuous function.



pdf and cdf for continuous RV



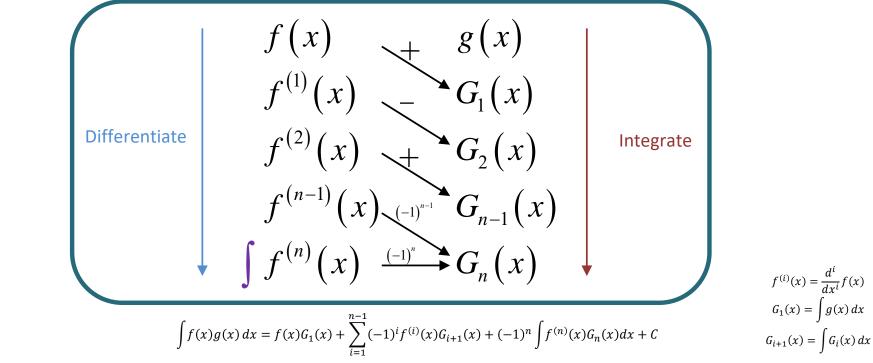
Finding Probabilities from CDF Definition: $F_X(x) \equiv P[X \leq x]$ For continuous RV, For any RV, • $P[X \leq b] = F_X(b)$ • $P[X \leq b] = F_X(b)$ $P[X < b] = F_X(b)$ $P[X < b] = F_X(b) - P[X = b]$ • $P[X > a] = 1 - F_x(a)$ • $P[X > a] = 1 - F_X(a)$ $P[X \ge a] = 1 - F_X(a) + P[X = a]$ $P[X \ge a] = 1 - F_X(a)$ • $P[a < X \le b] = F_X(b) - F_X(a)$ • $P[a < X \le b] = F_X(b) - F_X(a)$ $P[a < X < b] = F_{X}(b) - F_{X}(a)$ $P[a \le X < b] = F_X(b) - F_X(a)$ $P[a \le X \le b] = F_X(b) - F_X(a)$ • $P[X = a] = F_X(a) - F_X(a^{-})$ • P[X = a] = 0(amount of jump in the CDF (a) a) 23

Probability and Random Processes ECS 315

Asst. Prof. Dr. Prapun Suksompong prapun@siit.tu.ac.th 10.3 Expectation and Variance

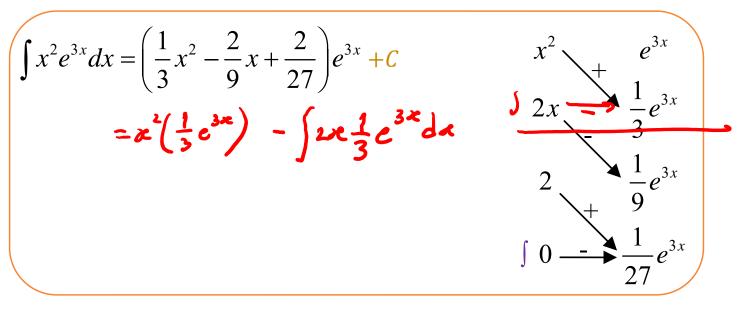
Integration by Parts

- A technique for simplifying integrals of the form $\int f(x)g(x) dx$
- **Tabular integration by parts**: A convenient method for organizing repeated application of integration by part:



[A.15]

Integration by Parts



$$\int (\sin x) e^x dx$$

= $(\sin x - \cos x) e^x - \int (\sin x) e^x dx$
= $\frac{1}{2} (\sin x - \cos x) e^x + C$

26

Chapter 9 vs. Section 10.3Discrete RVContinuous RV
$$\mathbb{E}X = \sum_{x} xp_{X}(x)$$
 $\mathbb{E}X = \int_{-\infty}^{\infty} xf_{X}(x)dx$ $\mathbb{E}[g(X)] = \sum_{x} g(x)p_{X}(x)$ $\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x)f_{X}(x)dx$ $\mathbb{E}[X^{2}] = \sum_{x} x^{2}p_{X}(x)$ $\mathbb{E}[x^{2}] = \int_{-\infty}^{\infty} x^{2}f_{X}(x)dx$ $\operatorname{Var}[X] = \mathbb{E}[(X - \mathbb{E}X)^{2}] = \mathbb{E}[X^{2}] - (\mathbb{E}X)^{2}$

Probability and Random Processes ECS 315

Asst. Prof. Dr. Prapun Suksompong prapun@siit.tu.ac.th 10.4 Families of Continuous Random Variables

Johann Carl Friedrich Gauss



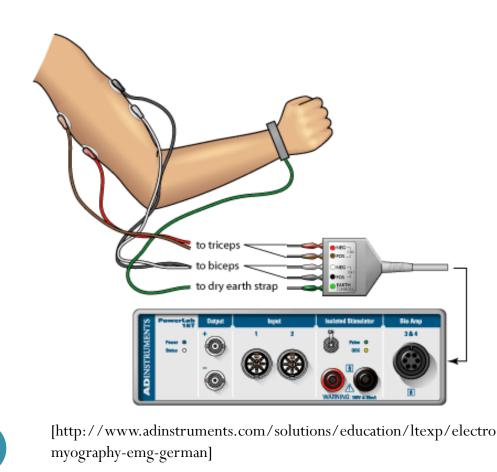
German 10-Deutsche Mark Banknote (1993; discontinued)

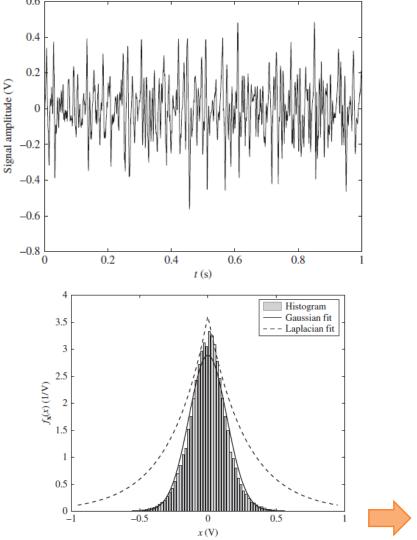
- 1777 1855
- A German mathematician



Ex. Muscle Activity

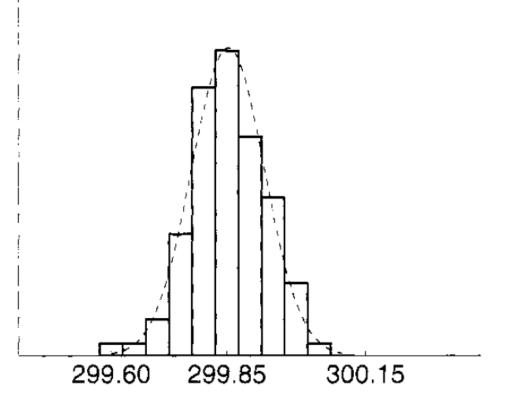
• Look at electrical activity of skeletal muscle by recording a human electromyogram (EMG).





Ex. Measuring the speed of light

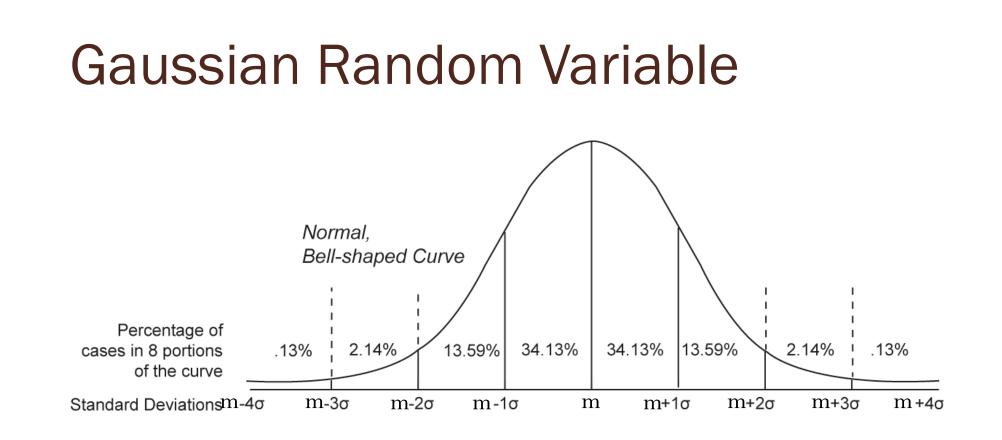
100 measurements of the speed of light (×1,000 km/second), conducted by Albert Abraham Michelson in 1879.

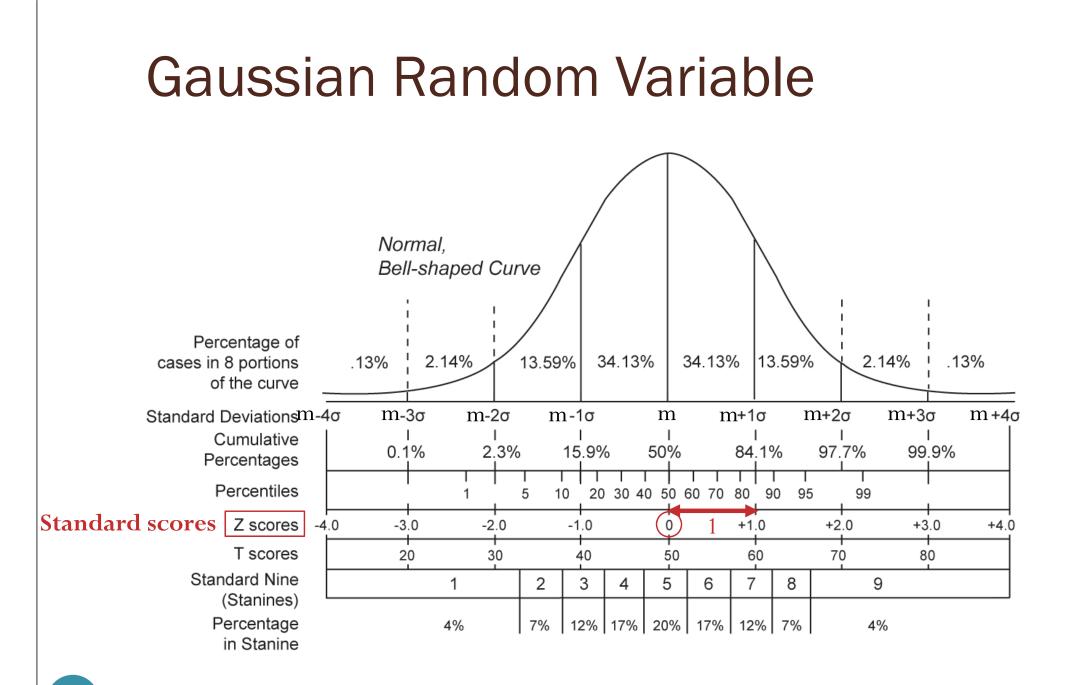


Expected Value and Variance

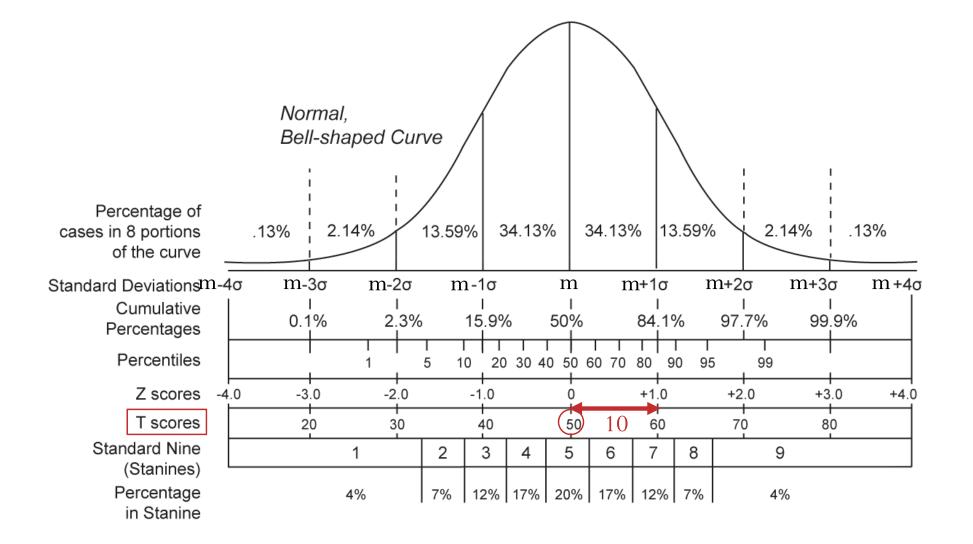
"Proof" by MATLAB's symbolic calculation

>> syms x >> syms m real >> syms sigma positive >> int(1/(sqrt(sym(2)*pi)*sigma)*exp(-(x-m)^2/(2*sigma^2)),x,-inf,inf) ans = 1 >> EX = $int(x/(sqrt(sym(2)*pi)*sigma)*exp(-(x-m)^2/(2*sigma^2)), x, -inf, inf)$ EX =m >> EX2 = $int(x^2/(sqrt(sym(2)*pi)*sigma)*exp(-(x-m)^2/(2*sigma^2)), x, -inf, inf)$ EX2 =-(2^(1/2)*(limit(- x*sigma^2*exp((x*m)/sigma^2 - m^2/(2*sigma^2) - x^2/(2*sigma^2)) - m*sigma^2*exp((x*m)/sigma^2 - m^2/(2*sigma^2) - x^2/(2*sigma^2)) -(2^(1/2)*pi^(1/2)*sigma*erfi((2^(1/2)*(x - m)*i)/(2*sigma))*(m^2 + sigma^2)*i)/2, x == -Inf) - limit(- x*sigma^2*exp((x*m)/sigma^2 - m^2/(2*sigma^2) x²/(2*sigma²)) - m*sigma²*exp((x*m)/sigma² - m²/(2*sigma²) - x²/(2*sigma²)) - (2^(1/2)*pi^(1/2)*sigma*erfi((2^(1/2)*(x - m)*i)/(2*sigma))*(m² + sigma^2)*i)/2, x == Inf)))/(2*pi^(1/2)*sigma) >> EX2 = simplify(EX2) EX2 = $m^2 + sigma^2$ >> VarX = EX2 - (EX)^2 VarX = sigma^2



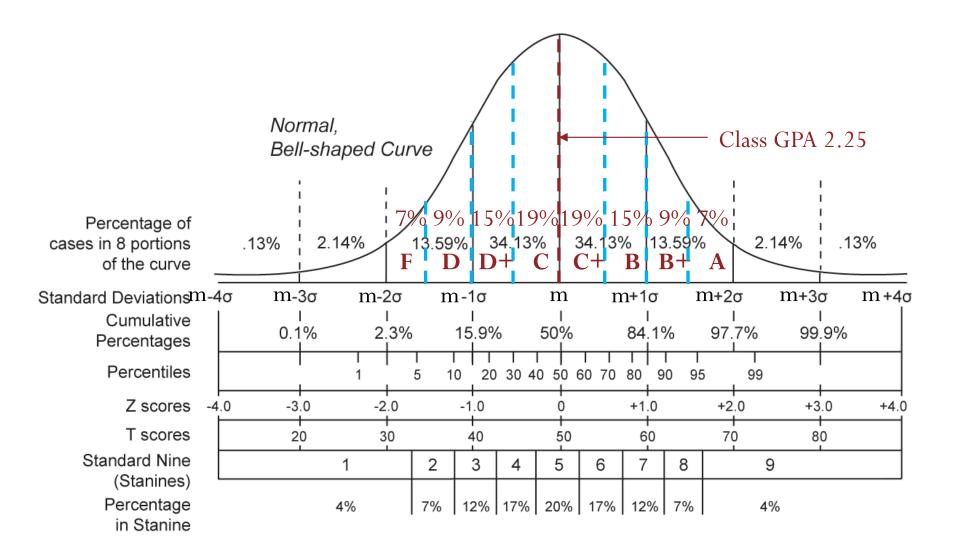


Gaussian Random Variable



[Wikipedia.org]

SIIT Grading Scheme (Option 3)



[Wikipedia.org]

From the News

Higgs boson-like particle discovery claimed at LHC

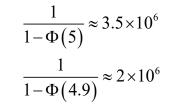
COMMENTS (1665)

By Paul Rincon

4 July 2012



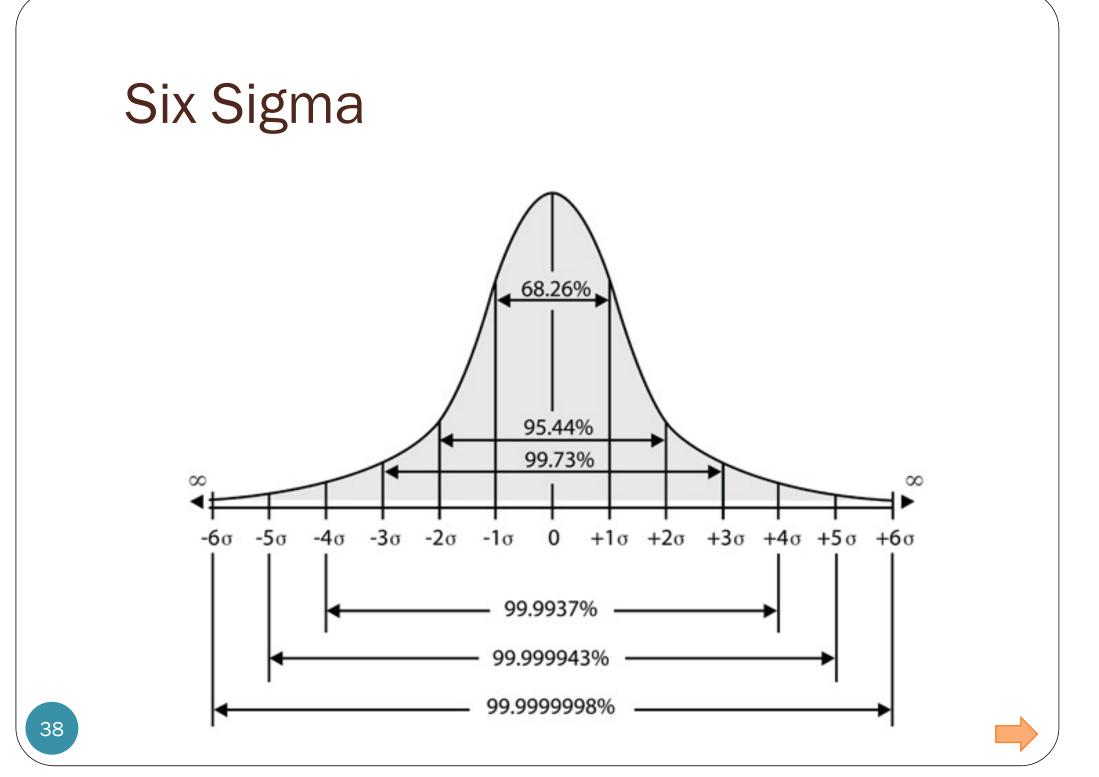




Particle physics has an accepted definition for a **discovery**: a "fivesigma" (or five standard-deviation) level of certainty **The number of sigmas measures how unlikely it is to get a certain experimental result as a matter of chance rather than due to a real effect**

They claimed that by combining two data sets, they had attained a confidence level just at the "five-sigma" point about a **one-in-3.5 million chance** that the signal they see would appear if there were no Higgs particle.

However, a full combination of the CMS data brings that number just back to **4.9 sigma** - a one-in-two million chance.



Six Sigma

- If you **manufacture** something that has a normal distribution and get an observation outside six σ of μ , you have either seen something extremely unlikely or there is something wrong with your manufacturing process. You'd better look it over.
- This approach is an example of **statistical quality control**, which has been used extensively and saved companies a lot of money in the last couple of decades.
- The term **Six Sigma**, a registered trademark of **Motorola**, has evolved to denote a methodology to monitor, control, and improve products and processes.
- There are Six Sigma societies, institutes, and conferences.
- Whatever Six Sigma has grown into, it all started with considerations regarding the normal distribution.

Six Sig	na $\int_{-6\sigma}^{\infty} -5\sigma -4\sigma -3\sigma -2}$	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$
Range around μ	Percentage of products in conformance	Percentage of nonconforming products
-1σ to $+1\sigma$	68.26	31.74
10 00 110	95.46	4.54
-2σ to $+2\sigma$		
	99.73	0.27
-3σ to $+3\sigma$	99.73 99.9937	$\begin{array}{c} 0.27 \\ 0.0063 \end{array}$
$\begin{array}{l} -2\sigma \text{ to } +2\sigma \\ -3\sigma \text{ to } +3\sigma \\ -4\sigma \text{ to } +4\sigma \\ -5\sigma \text{ to } +5\sigma \end{array}$		

Probabilities involving Gaussian RV

• There is no closed-form simplification for

$$\int_{a}^{b} \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^{2}} dx. \quad (\text{except for some special cases})$$

• We have a table which gives the cdf of a standard Gaussian RV:

 $\Phi(z) \equiv F_Z(z)$ when $Z \sim \mathcal{N}(0,1)$.

- The Φ table gives $\Phi(z)$ for $z \in [0,3)$.
- Can use the property

$$\Phi(-z) = 1 - \Phi(z)$$

to work with z < 0

Probabilities involving Gaussian RV

• There is no closed-form simplification for (except for some special cases)

$$\int_{a}^{b} \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^{2}} dx.$$

- We have a table which gives the cdf of a standard Gaussian RV: $\Phi(z) \equiv F_Z(z)$ when $Z \sim \mathcal{N}(0,1)$.
 - The Φ table gives $\Phi(z)$ for $z \in [0,3)$.
 - The Q table gives $Q(z) = 1 \Phi(z)$ for $z \in [3,5)$.
 - Can use the property $\Phi(-z) = 1 \Phi(z)$ to work with z < 0
- For $X \sim \mathcal{N}(m, \sigma^2)$,
 - $P[X \le b] = P[X \le b] = F_X(b) = \Phi\left(\frac{b-m}{\sigma}\right)$
 - $P[X > a] = P[X \ge a] = 1 F_X(a) = 1 \Phi\left(\frac{a m}{\sigma}\right)$
 - $P[a < X \le b] = P[a < X < b] = P[a \le X < b] = P[a \le X < b] = P[a \le X \le b]$ = $F_X(b) - F_X(a) = \Phi\left(\frac{b-m}{\sigma}\right) - \Phi\left(\frac{a-m}{\sigma}\right)$

More on Gaussian RVs...

volume 150

HANDBOOK OF THE NORMAL DISTRIBUTION

CO STATISTICS: textbooks and monographs

Second Edition, Revised and Expanded

JAGDISH K. PATEL CAMPBELL B. READ

Probability Distributions Involving Gaussian Random Variables

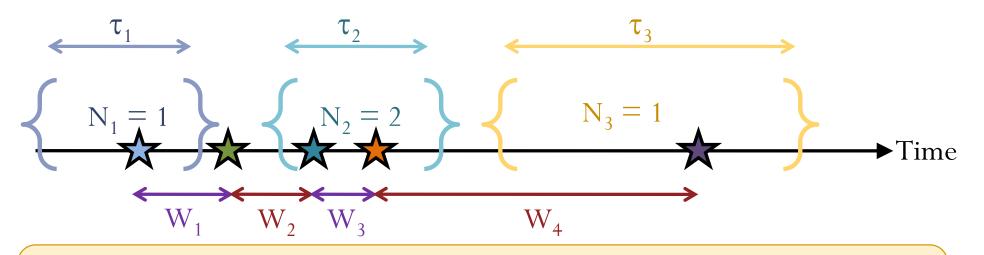
A Handbook for Engineers, Scientists and Mathematicians

Marvin K. Simon

Deringer

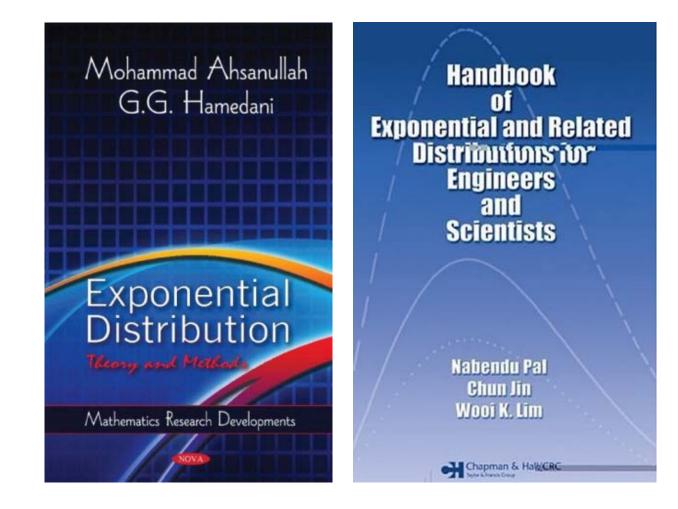
Poisson Process

The number of arrivals N_1, N_2, N_3, \ldots during non-overlapping time intervals are independent Poisson random variables with mean = $\lambda \times$ the length of the corresponding interval.



The lengths of time between adjacent arrivals W_1, W_2, W_3, \ldots are i.i.d. exponential random variables with mean $1/\lambda$.

More on Exponential RV ...



References

- From Discrete to Continuous Random Variables: [Y&G]
 Sections 3.0 to 3.1
- PDF and CDF: [Y&G] Sections 3.1 to 3.2
- Expectation and Variance: [Y&G] Section 3.3
- Families of Continuous Random Variables: [Y&G] Sections 3.4 to 3.5

Course Outline

The following is a tentative list of topics with their corresponding chapters from the text Yates and Goodman. Each topic spans approximately one week.

1.	Introduction, Set Theory, Classical Probability	[1]
2.	Combinatorics: Four Principles and Four Kinds of Counting Problems	[1]
3.	Probability Foundations	[1]
4.	Event-based Conditional Probability	[1]
5.	Event-based Independence	[1]
6.	Random variables, Support, Probability Distribution	[2]
7.	MIDTERM: 4 Oct 2018 TIME 09:00 - 11:00	
8.	Discrete Random Variables	[2]
9.	Families of Discrete Random Variables and Introduction to Poisson	[2,10]
	Processes	
10.	Real-Valued Functions of a Random Variable	[2]
11.	Expectation, Moment, Variance, Standard Deviation	[2]
12.	Continuous Random Variables	[3]
13.	Families of Continuous Random Variables and Introduction to	[3,10]
	Poisson Processes	

Excercise 15 Solution [Posted @ 4:30PM on I

- Excercise 16 Solution [Posted @ 3PM on Nov
- Slides [Posted @ 4:30PM on Nov 6]
- Part IV: Continuous Random Variables
 - Chapter 10 [Posted @ 10AM on Nov 5]
 - Annotated notes for Sections 10.1-10.3 [Pos
 - References
 - From Discrete to Continuous Random
 - PDF and CDF: [Y&G] Sections 3.1 to 3.2
 - Expectation and Variance: [Y&G] Sectio
 - Families of Continuous Random Variab
- Part V: Multiple Random Variables

Probability and Random Processes ECS 315

Asst. Prof. Dr. Prapun Suksompong prapun@siit.tu.ac.th 10.5

Review: Function of discrete RV

Example 9.16. Let

$$p_X(x) = \begin{cases} \underbrace{1 \atop x^2}, & x = \pm 1, \pm 2\\ 0, & \text{otherwise} \end{cases}$$

and

48

$$Y = X^4$$

Find $p_Y(y)$ and then calculate $\mathbb{E}Y := \sum_{y} y p_y (y)$

Step 1: Find c

$$\sum_{x} p_{x}(x) = 1$$
Note that $Y = X^{4}$

$$P_{Y}(Y) = \begin{cases} 1/5, & y = 1, \\ 4/5, & y = 1L, \\ 0, & 0 \text{ thermise.} \end{cases}$$

$$\frac{1}{c} \left(1^{\frac{3}{2}} + 2^{\frac{3}{2}} + (-1)^{\frac{3}{2}} + (-2)^{\frac{3}{2}}\right) = 1$$

$$p_{x}(x) = x \qquad Y$$

$$\frac{1}{1/10} \qquad 1 \qquad 1^{\frac{4}{2}} = 1$$

$$C = 10, \qquad 1/10 \qquad -1 \qquad (-1)^{\frac{4}{2}} = 1$$

$$\frac{4/40}{2} \qquad 2 \qquad 2^{\frac{4}{2}} = 11$$

$$= 1 \times \frac{1}{5} \qquad + \qquad 16 \times \frac{4}{5}$$

$$P[Y = 1] = P[X = 1] + P[X = -1] = \frac{2}{10} = \frac{1}{5} \qquad = \frac{65}{5} = 13$$

$$P[Y = 1L] = P[X = 2] + P[X = -2] = \frac{8}{10} = \frac{4}{5}$$

References

- From Discrete to Continuous Random Variables: [Y&G] Sections 3.0 to 3.1
- PDF and CDF: [Y&G] Sections 3.1 to 3.2
- Expectation and Variance: [Y&G] Section 3.3
- Families of Continuous Random Variables: [Y&G] Sections
 3.4 to 3.5
- SISO: [Y&G] Section 3.7; [Z&T] Section 5.2.5