VANDERBILT
School *of* Engineering

CE 3300-01 – RISK, RELIABILITY, AND RESILIENCE ENGINEERING

Multiple Random Variables

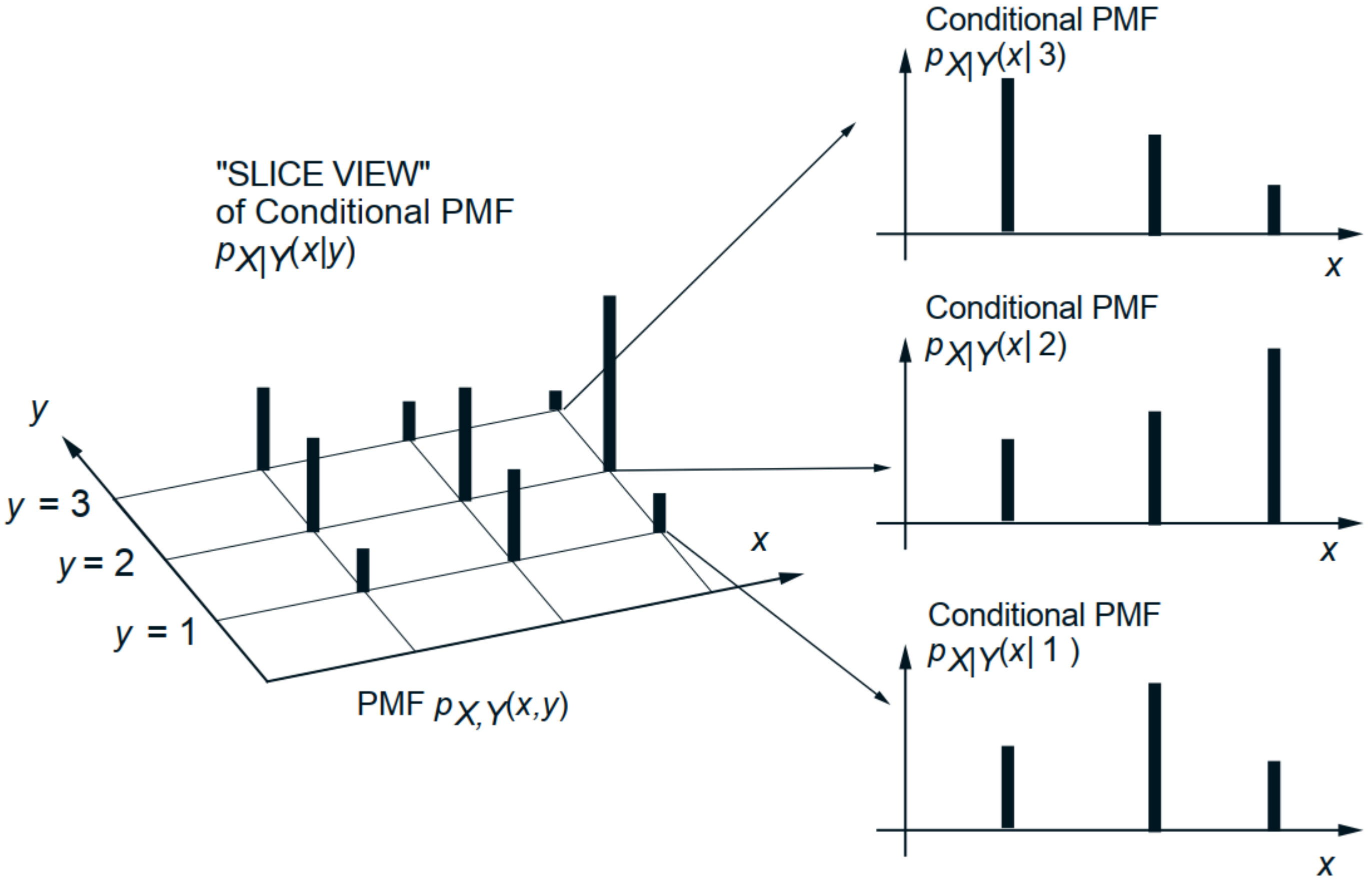
Book Reference: Chapters 3-4

TA: Mohamad Kazma

February 18th-19th, 2025

Instructor: Dr. Hiba Baroud

Today: Multiple Random Variables



Recap: Standard Normal Distribution

Let X be a normal random variable with mean μ and variance σ^2.

We "standardize" X by defining a new random variable Y given by

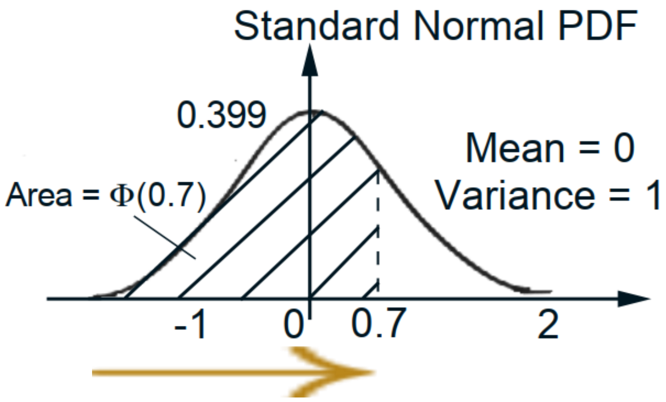
$$Y = \frac{X-\mu}{\sigma} \qquad Y \sim \mathcal{N}\left(\mu = 0, \sigma^2 = 1\right)$$

A normal random variable Y with zero mean and unit variance is said to be a standard normal. Its CDF is denoted by Φ,

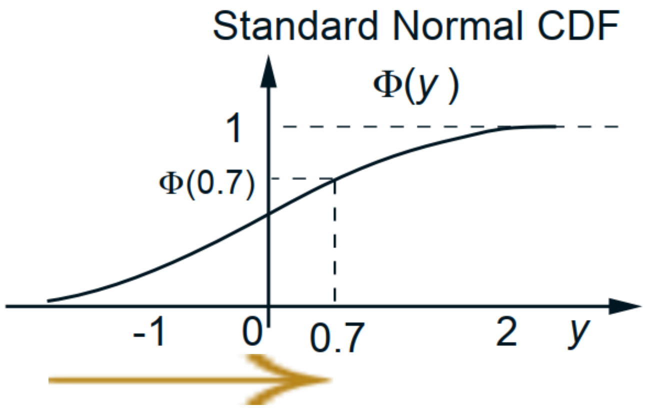
$$\Phi(y) = \mathbf{P}(Y \leq y) = \mathbf{P}(Y < y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-t^2/2} dt$$

Complement

$$\mathbf{P}(Y > y) = 1 - \mathbf{P}(Y \leq y)$$



Standard Normal PDF



Standard Normal CDF

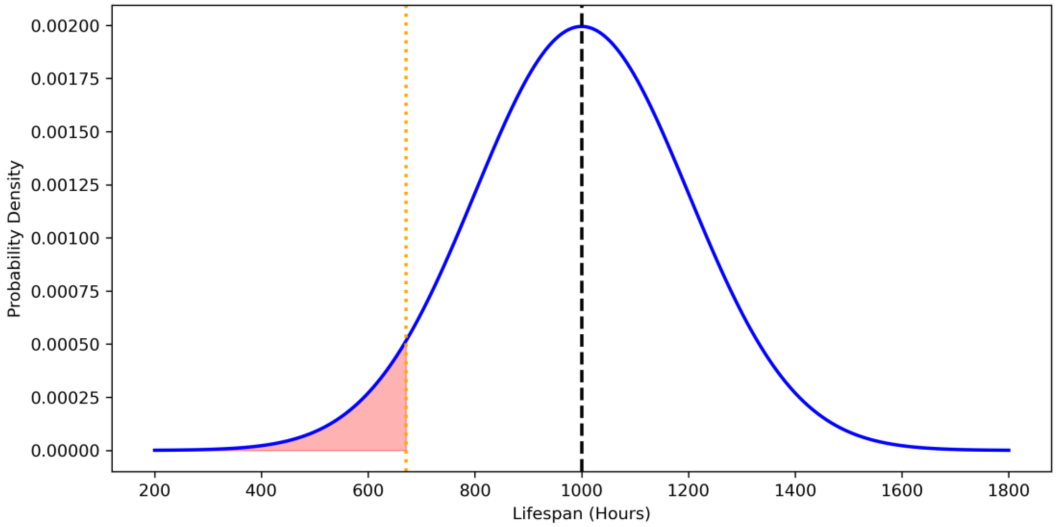
Negative Values

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

Recap: Example 4 – Lifespan of Lightbulbs

b. **What should the warranty period be if the company wants to cover only the bottom 5% of bulbs?**

We want to find the value of X where 5% of the data is below it.



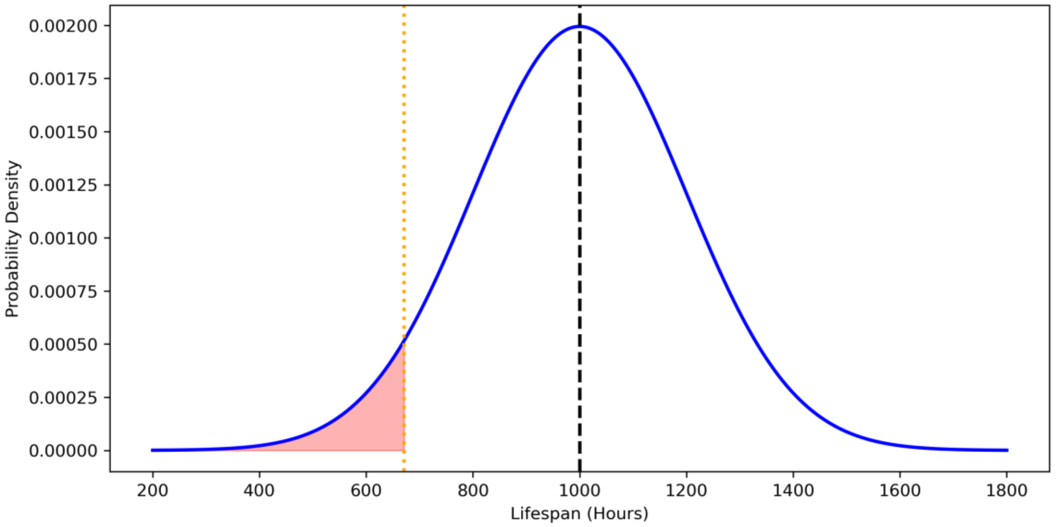
Let $X \sim \mathcal{N}\left(\mu = 1000, \sigma^2 = 200^2\right)$

Then, $Z \sim \mathcal{N}\left(\mu = 0, \sigma^2 = 1\right)$

b. We are looking for the value of $P(Z < z) = 0.05$
From the Z-table, what is the Z-score corresponding to $P(Z) = 0.05$

Recap: Example 4 – Lifespan of Lightbulbs

b. **What should the warranty period be if the company wants to cover only the bottom 5% of bulbs?**



If $P(Z < z)$ is to be 0.05
then, we can say given the symmetry of the normal distribution that $P(Z > z)$ is to be 0.95

Then to find z for the probability of 0.95 we refer to Appendix A.

From Appendix A, for $\Phi(z) = 0.95$, we have $z = 1.645$

we now refer to the relation → $\Phi(-y) = 1 - \Phi(y)$

As such, $0.05 = 1 - \Phi(z) = 1 - \Phi(1.645) = \Phi(-1.645)$ $\boxed{\rightarrow z = -1.645}$

$$\boxed{\text{Then, } X = \mu + Z\sigma = 1000 + (-1.645)(200) = 670}$$

Multiple RVs – Learning Objectives

- Probability Distributions of Multiple RVs
  - **Describe** probability distributions of multiple RVs

  - **Calculate** probability of multiple probability distributions

  - **Solve** problems with both discrete and continuous R.V.

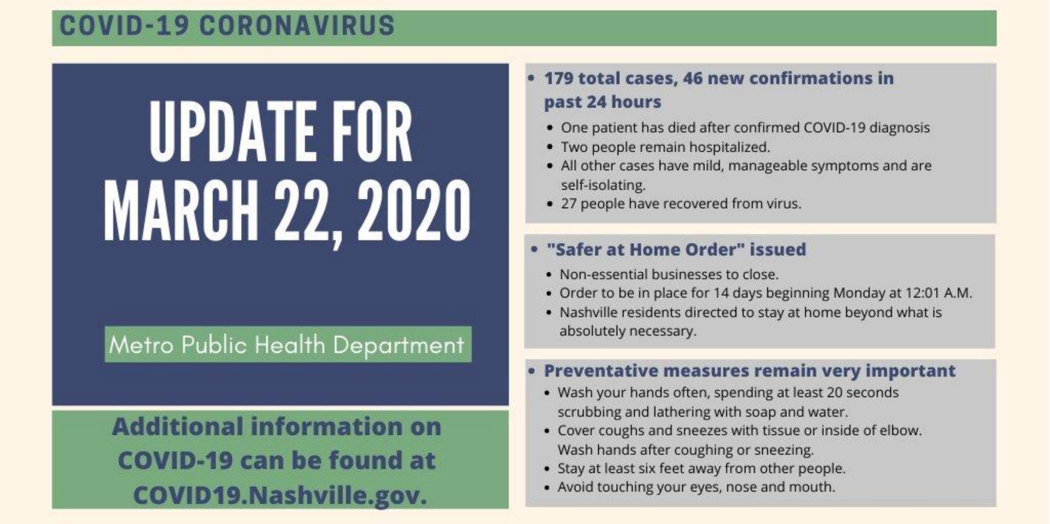
Univariate vs. Multivariate distributions

Interest in evaluating the probability of multiple events

- E.g., What is the joint probability of having an earthquake and a tsunami?

Multiple Hazard

- Wind and surge (correlated)
- Earthquake and flood (uncorrelated)
- Wind and earthquake (different time scales)
- Fire after earthquake (cascading effects)
- Tornado and pandemic (Nashville)



COVID-19 CORONAVIRUS

UPDATE FOR MARCH 22, 2020

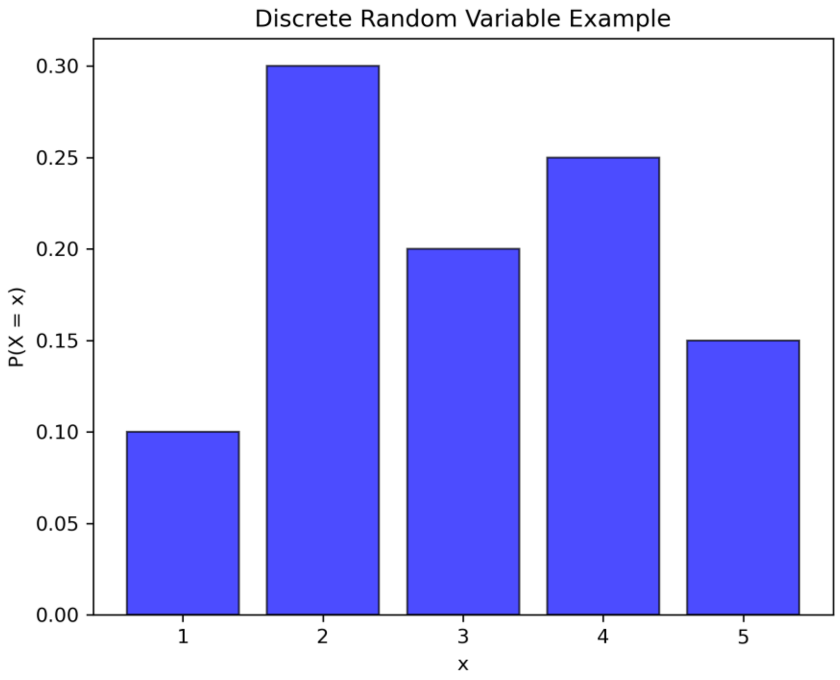
Metro Public Health Department

Additional information on COVID-19 can be found at COVID19.Nashville.gov.

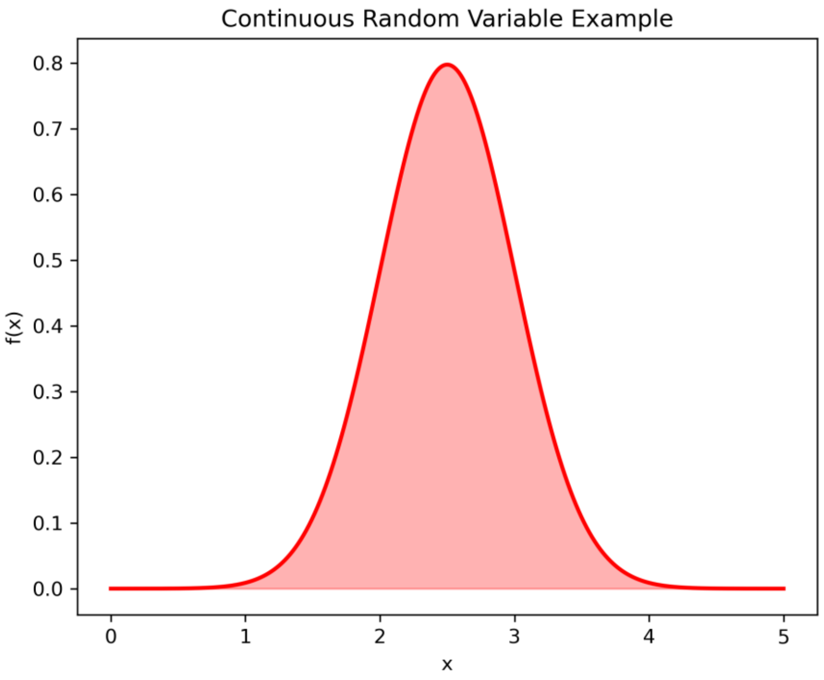
- **179 total cases, 46 new confirmations in past 24 hours**
  - One patient has died after confirmed COVID-19 diagnosis
  - Two people remain hospitalized.
  - All other cases have mild, manageable symptoms and are self-isolating.
  - 27 people have recovered from virus.
- **"Safer at Home Order" issued**
  - Non-essential businesses to close.
  - Order to be in place for 14 days beginning Monday at 12:01 A.M.
  - Nashville residents directed to stay at home beyond what is absolutely necessary.
- **Preventative measures remain very important**
  - Wash your hands often, spending at least 20 seconds scrubbing and lathering with soap and water.
  - Cover coughs and sneezes with tissue or inside of elbow. Wash hands after coughing or sneezing.
  - Stay at least six feet away from other people.
  - Avoid touching your eyes, nose and mouth.

Recall: Discrete Vs. Continuous RVs

- A RV is a variable whose value can take on one of several realizations such that each realization is associated with a likelihood of occurrence



$$P(X = x)$$



$$P(a < X \leq b) = \int_a^b f_X(x)dx$$

Joint PMF of Two Discrete RVs

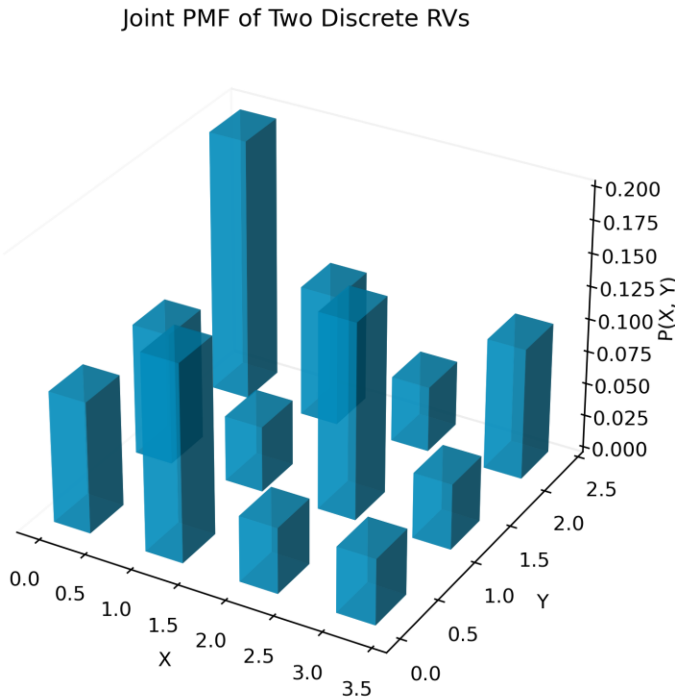
Joint PMF of two Discrete RVs

$$p_{X,Y}(x, y) = \mathbf{P}(X = x, Y = y) = \mathbf{P}(\{X = x\} \cap \{Y = y\})$$

Axioms

$$0 \leq p_{XY}(x, y) \leq 1$$

$$\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} p_{XY}(x_i, y_j) = 1$$



CDF and the Marginal Probability of Two Discrete RVs

CDF of two Discrete RVs

$$F_{XY}(x,y) = P(X \leq x_i \cap Y \leq y_i) = \sum\nolimits_{\forall x_i \leq x} \sum\nolimits_{\forall y_j \leq y} p_{XY}(x_i, y_j)$$

$$P(a \leq X \leq b \cap c \leq Y \leq d) = \sum\nolimits_{i=a}^{b} \sum\nolimits_{j=c}^{d} p_{XY}(x_i, y_j)$$

Marginal Probabilities

$$p_X(x) = \sum_{j=0}^{\infty} p_{XY}(x, y_j) \qquad p_Y(y) = \sum_{i=0}^{\infty} p_{XY}(x_i, y)$$

Recall: $F_X(x) = \sum\nolimits_{\text{all } x_i \leq x} P(X = x_i) = \sum\nolimits_{\text{all } x_i \leq x} p_X(x_i)$

Marginal Probability of Two Discrete RVs

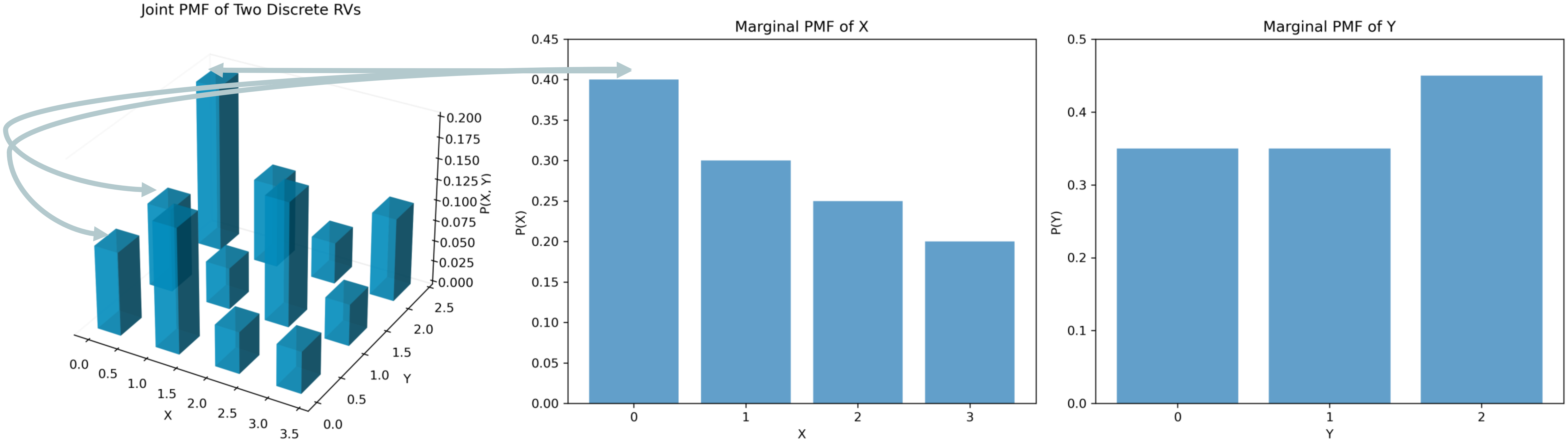
Marginal Probabilities

- There is no difference between the marginal probability that we discussed during our probability rules lectures for single random variables and the marginal PMF $p_X(x)$. They both give the probability that $X = x$.

- However, the word *marginal* implies that there is an additional random variable that is being considered.

- In general, if the joint PMF is known, it is always possible to get the marginal PMF's. However, if the marginal PMF's are known, it is usually not possible to get the joint PMF.

Marginal Probability of Two Discrete RVs



Conditional Probability of Two Discrete RVs

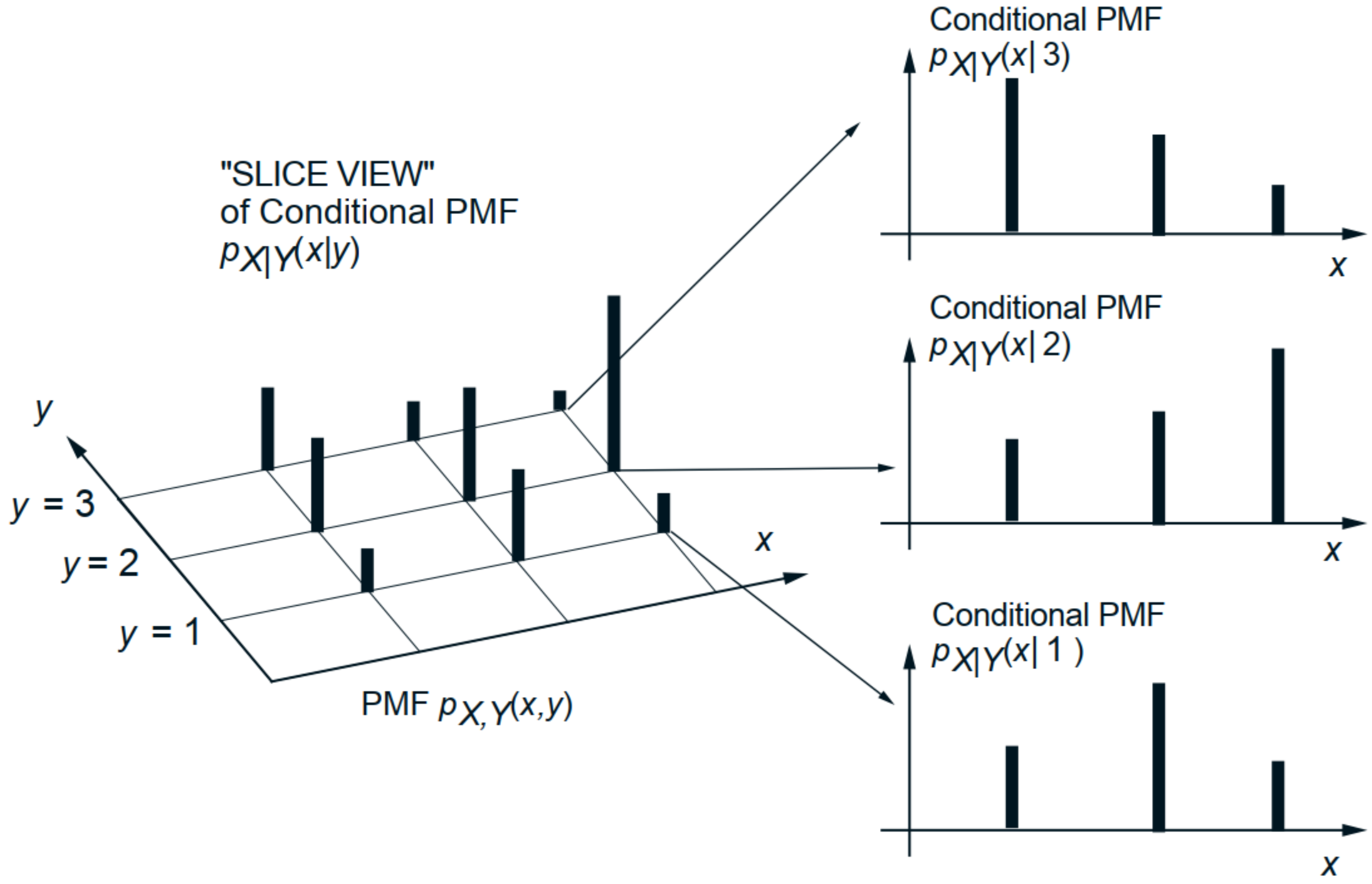
Conditional probability of two Discrete RVs

Conditional probability mass function (conditional PMF), $p_{X|Y}(x,y)$, is the probability mass function of X given that $Y = y$

$$p_{X|Y}(a,b) = \frac{P(X=a \cap Y=b)}{P(Y=b)} = \frac{p_{XY}(a,b)}{p_Y(b)}$$

Recall: $P(A \mid B) = \frac{P(A \cap B)}{P(B)}$

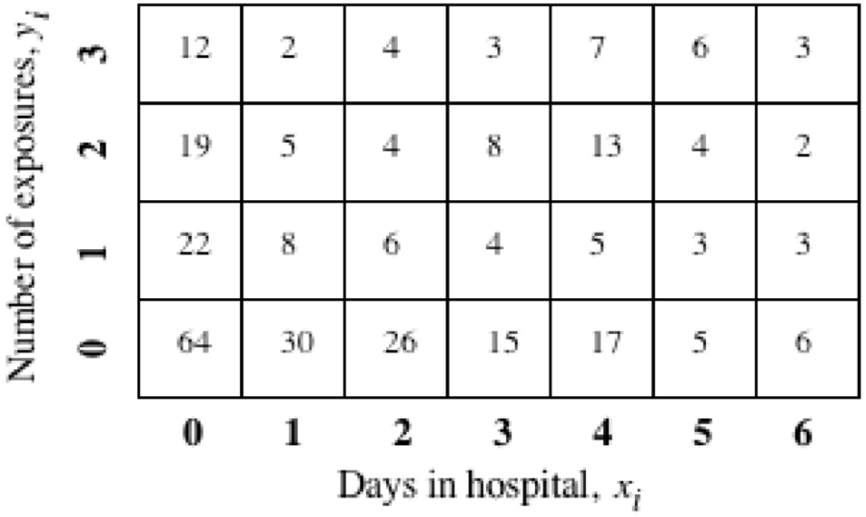
Conditional Probability of Two Discrete RVs



Example 1 – Factory Leak

In a factory, there is occasionally a small leak of foul-smelling gas.
The factory owners claim that the gas has no effect on health.
However, the workers complain, so the owners call in an environmental specialist to study the situation.
Upon inspection, the environmental specialist that employees have had variable exposure to the gas. Further, some of the employees have spent time in the hospital recently.

The specialist gathers data. For each of the **306 workers**, the **number of days spent in a hospital** for respiratory problems, x_i, and **the number of exposures**, y_i, are recorded.

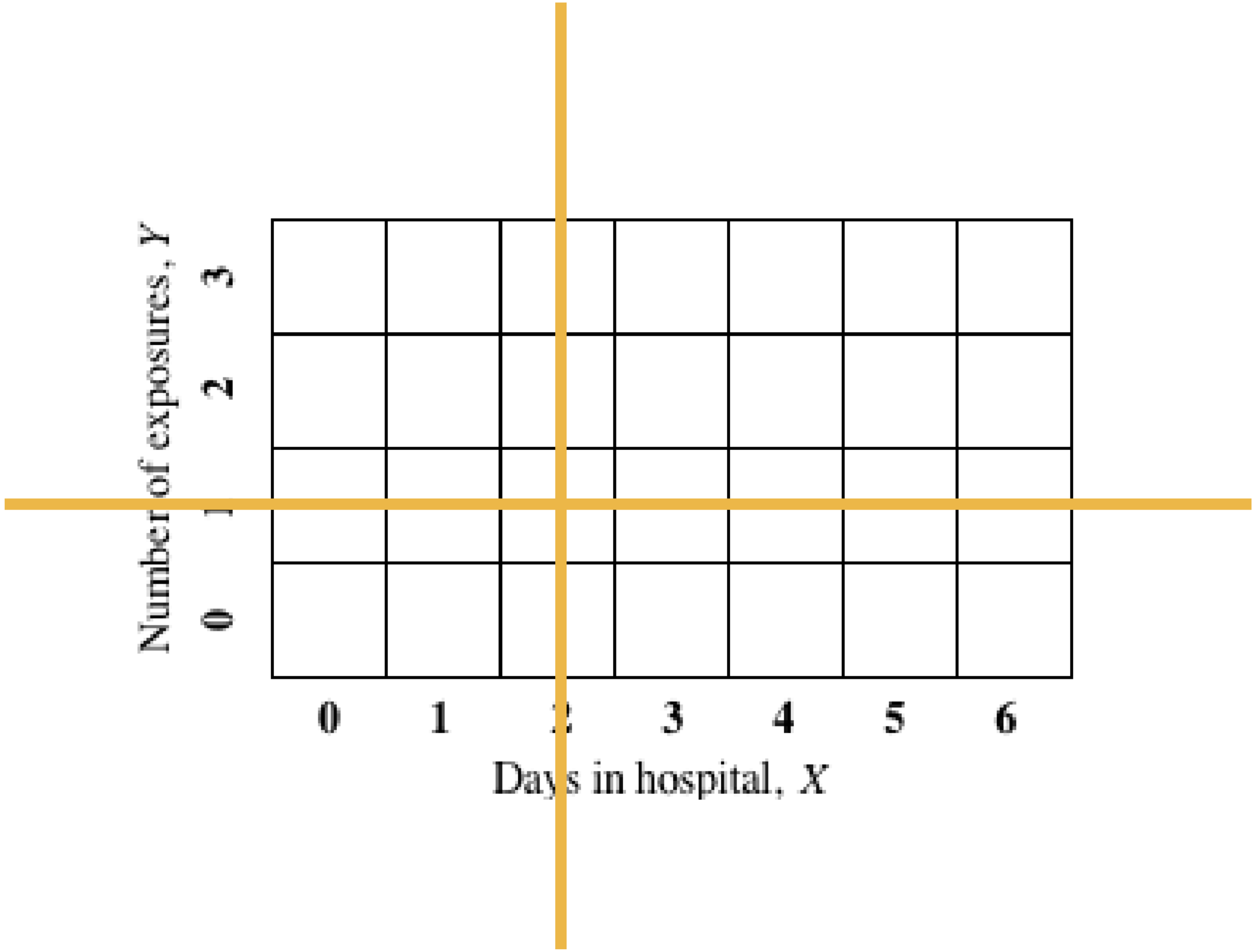


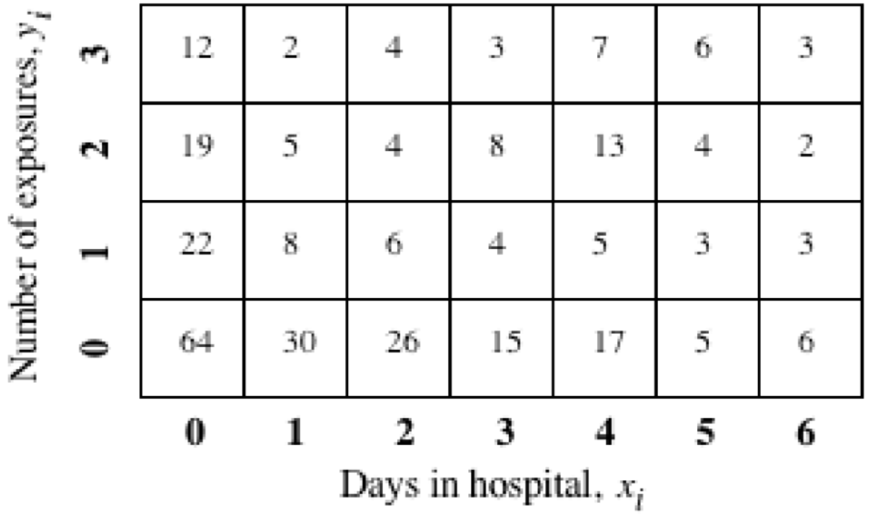
Number of exposures, y_i

3	12	2	4	3	7	6	3
2	19	5	4	8	13	4	2
1	22	8	6	4	5	3	3
0	64	30	26	15	17	5	6
	0	**1**	**2**	**3**	**4**	**5**	**6**

Days in hospital, x_i

Example 1 – Continued





Number of exposures, y_i

3	12	2	4	3	7	6	3
2	19	5	4	8	13	4	2
1	22	8	6	4	5	3	3
0	64	30	26	15	17	5	6
	0	**1**	**2**	**3**	**4**	**5**	**6**

Days in hospital, x_i

$$p_{XY}(x = 2, y = 1) = \frac{6}{306}$$

Example 1 – Continued

Joint probability mass function $f_{XY}(x, y)$
Calculate the marginal PMF $P_X(X = 5)$

Number of exposures, Y

3	.0392	.0065	.0131	.0098	.0229	.0196	.0098
2	.0621	.0163	.0131	.0261	.0425	.0131	.0065
1	.0719	.0261	.0196	.0131	.0163	.0098	.0098
0	.2092	.0980	.0850	.0490	.0556	.0163	.0196
	0	**1**	**2**	**3**	**4**	**5**	**6**

Days in hospital, X

$$
\begin{aligned}
p_X(x = 5) &= \sum_{i=0}^{3} p_{XY}\left(x = 5, y_i\right) \\
&= p_{XY}(5, 0) + p_{XY}(5, 1) + p_{XY}(5, 2) + p_{XY}(5, 3) \\
&= 0.0588
\end{aligned}
$$

Example 1 – Continued

What is the probability mass function (PMF) of exposure if a worker spent 5 days at the hospital?

Number of exposures, Y

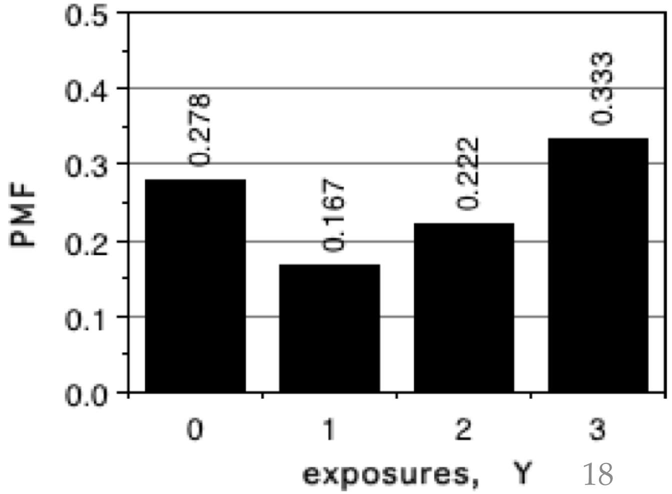
3	.0392	.0065	.0131	.0098	.0229	.0196	.0098
2	.0621	.0163	.0131	.0261	.0425	.0131	.0065
1	.0719	.0261	.0196	.0131	.0163	.0098	.0098
0	.2092	.0980	.0850	.0490	.0556	.0163	.0196
	0	**1**	**2**	**3**	**4**	**5**	**6**

Days in hospital, X

$$p_{Y|X}(y,5) = ?$$

$$p_{Y|X}(0,5) = \frac{p_{XY}(5,0)}{p_X(5)} = \frac{0.0163}{0.0588} = 0.277$$

$$p_{Y|X}(1,5) = \frac{p_{XY}(5,1)}{p_X(5)} = \frac{0.0098}{0.0588} = 0.167$$



Example 1 – Continued

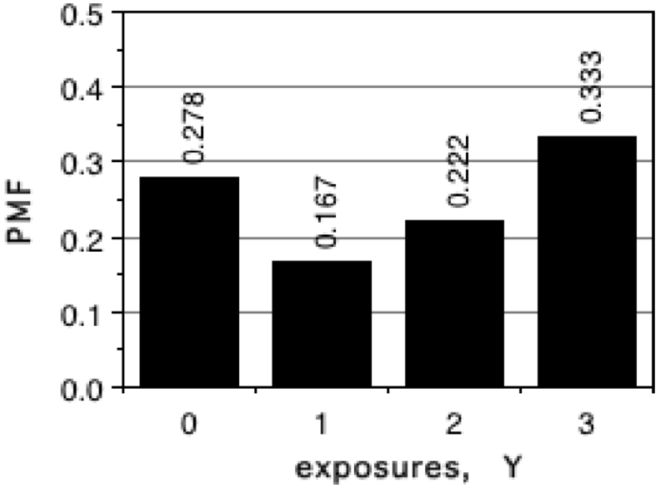
How is the conditional PMF different from the marginal PMF of Y?

Number of exposures, Y

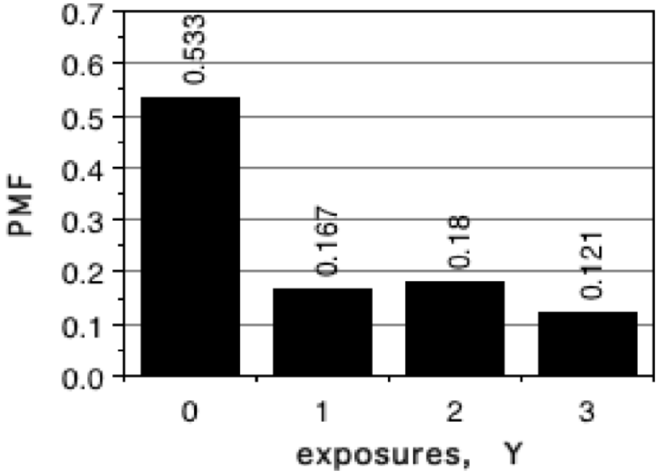
3	.0392	.0065	.0131	.0098	.0229	.0196	.0098
2	.0621	.0163	.0131	.0261	.0425	.0131	.0065
1	.0719	.0261	.0196	.0131	.0163	.0098	.0098
0	.2092	.0980	.0850	.0490	.0556	.0163	.0196
	0	**1**	**2**	**3**	**4**	**5**	**6**

Days in hospital, X

$$p_{y|x}(y, x = 5)$$



$$p_Y$$



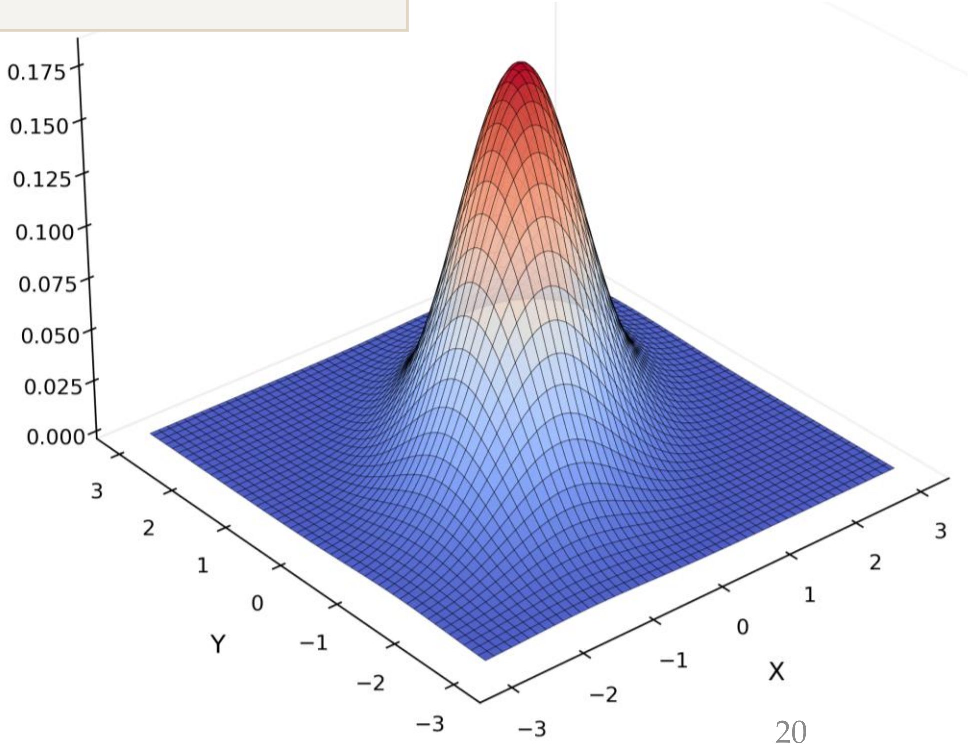
Joint PDF of Two Continuous RVs

Joint PDF of two Continuous RVs

$$\int_a^b \int_c^d f_{XY}(x,y)dydx = P(a < X < b, c < Y < d)$$

Axioms

$$\iint f_{XY}(x,y)dxdy = 1$$



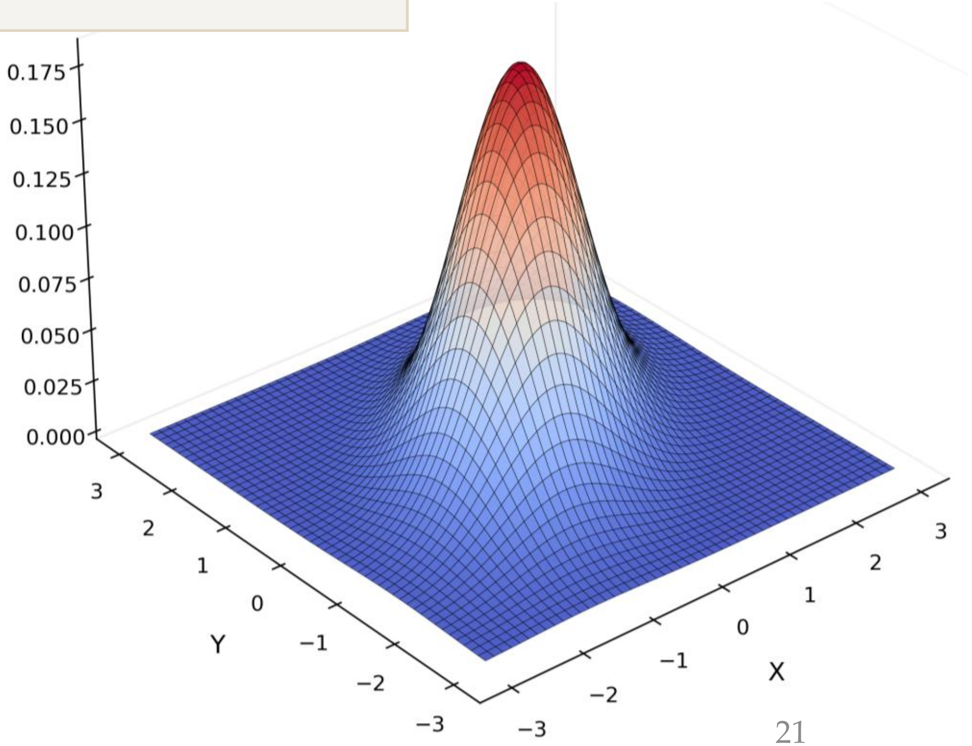
Joint PDF of Two Continuous RVs

Joint CDF of two Continuous RVs

$$F_{X,Y}(x,y) = \mathbf{P}(X \leq x, Y \leq y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(s,t)dsdt.$$

The PDF can be recovered from the PDF by differentiating:

$$f_{X,Y}(x,y) = \frac{\partial^2 F_{X,Y}}{\partial x \partial y}(x,y)$$



Conditional and Marginal PDF of Two Continuous RVs

Conditional PDF

$$\mathbf{P}(X \in A \mid Y = y) = \int_A f_{X|Y}(x \mid y)dx$$

↳ $$f_{X|Y}(x \mid y) = \frac{f_{XY}(x,y)}{f_Y(y)}$$

Marginal PDFs

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)dy \qquad f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)dx$$

Example 2

X & Y are jointly continuous R.V. with joint pdf:

$$f_{XY}(x,y) = \begin{cases} 10x^2y & 0 \leq y \leq x \leq 1 \\ 0 & Otherwise \end{cases}$$

a. Find the marginal pdfs

Range: $R_X = R_Y = [0,1]$

$$f_X(x) = \int_0^x 10x^2 y dy = 10x^2 \left(\frac{y^2}{2} \Big|_0^x \right) = 5x^4$$

$$f_X(x) = \begin{cases} 5x^4 & 0 \leq x \leq 1 \\ 0 & \text{Otherwise} \end{cases}$$

Example 2

X & Y are jointly continuous R.V. with joint pdf:

$$f_{XY}(x,y) = \begin{cases} 10x^2y & 0 \leq y \leq x \leq 1 \\ 0 & Otherwise \end{cases}$$

a. Find the marginal pdfs

Range: $R_X = R_Y = [0,1]$

$$f_Y(y) = \int_y^1 10x^2y\,dx = y \times \left. \frac{10x^3}{3} \right|_y^1 = \frac{10}{3} y \left(1 - y^3\right)$$

$$f_Y(y) = \begin{cases} \frac{10}{3} y \left(1 - y^3\right) & 0 \leq y \leq 1 \\ 0 & \text{Otherwise} \end{cases}$$

Example 2

X & Y are jointly continuous R.V. with joint pdf:

$$f_{XY}(x,y) = \begin{cases} 10x^2y & 0 \le y \le x \le 1 \\ 0 & Otherwise \end{cases}$$

b. Find $P\left(Y \le \frac{x}{2}\right)$

$$P\left(Y \le \frac{X}{2}\right) = \int_0^1 \int_0^{\frac{x}{2}} f_{XY}(x,y)dydx = \int_0^1 \int_0^{\frac{x}{2}} 10x^2ydydx = \int_0^1 \frac{5}{4}x^4dx = \frac{1}{4}$$

Sum of n Independent and Identically (iid) Distributed RVs

Independent and Identically Distributed (IID) random variables share the same probability distribution and are statistically independent.

$$Y = \sum_{i=1}^{n} [X_i] = X_1 + X_2 + \cdots + X_n$$

Mean and Variance

$$E[Y] = \sum_{i=1}^{n} E\left[X_i\right]$$

$$\operatorname{Var}[Y] = \sum_{i=1}^{n} \operatorname{Var}\left[X_i\right]$$

Sum of n iid RVs

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

P(X)	**P(Y)**
Normal (μ, σ)	Normal $(n\mu, \sqrt{n}\sigma)$
Binomial (n, p)	Binomial $(\sum n, p)$
Bernoulli (p)	Binomial (n, p)
Poisson(λ)	Poisson $(\sum \lambda)$
Exponential (λ)	Gamma(n, λ)

A6 Team Distributions

1. Geometric distribution --- Probability Princesses (Rachel, Jackie, Veronica, Maddie, Drew)

2. Hypergeometric distribution --- The Blueprint Bandits (Logan, Zach, Winnie, Elise, Nevaeh)

3. Gamma distribution --- Standard Deviants (Sydney, Ella, Yidi, Diego, Natalie)

4. Negative binomial distribution --- VennDyDiagram (Jacob, Joshua, Alyssa, Eden, Sharon)

5. Uniform distribution --- You Miss Every Shot You Don't Take (Brandon, Sarah, Caroline, Taha, Fiona, Ben)

Note: Make sure you join one of the teams. Four teams of 5 members and one team of 6 members.