1179: Probability
Lecture 22 — Bivariate Normal and MGF

Ping-Chun Hsieh (謝秉均)

December 3, 2021

Clarification: Cauchy-Schwarz Inequality

- **Cauchy-Schwarz Inequality**: Let X, Y be two random variables. Then, we have

$$E[X^2] \cdot E[Y^2] \geq (E[XY])^2$$

- Question: Under what condition do we have "="?

There exists some $t \in \mathbb{R}$ such that

$$P(\{\omega : Y(\omega) = t \cdot X(\omega)\}) = 1 \iff E[X^2] \cdot E[Y^2] = (E[XY])^2$$

$$\Longleftarrow$$

Quick Overview

- Given 2 random variables X, Y: what have we learned so far?

1. Joint CDF
2. Marginal CDF
3. Joint PMF / PDF
4. Marginal PMF / PDF
5. Independence
6. Conditional distribution
7. Expected value involving both X, Y
8. Bivariate normal
9. Distribution of $X + Y$
10. Covariance and correlation

This Lecture

1. Construction of Bivariate Normal

2. Moment Generating Functions

- Reading material: Chapter 10.5 and 11.1

Review: Construction of Bivariate Normal R.V.

- **Idea:** Let Z, W be 2 independent standard normal r.v.s and $\rho \in [-1, 1]$. Define two random variables

$$\begin{cases} X_1 = \sigma_1 Z + \mu_1 \\ X_2 = \sigma_2\left(\rho Z + \sqrt{1-\rho^2}\,W\right) + \mu_2 \end{cases}$$

$X_1 \sim N(\mu_1, \sigma_1^2)$

$X_2 \sim N(\mu_2, \sigma_2^2)$

- **Question:** Is it possible to find the joint PDF of X_1, X_2?

$$f_{X_1X_2}(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left[- \frac{\left(\frac{(x_1-\mu_1)^2}{\sigma_1^2} - 2\rho \frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2} \right)}{2(1-\rho^2)} \right]$$

Bivariate Normal R.V.s (Formally)

- **Bivariate Normal**: X_1 and X_2 are said to be bivariate normal random variables if the joint PDF of X_1, X_2 is

$$f_{X_1X_2}(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left[-\frac{\left(\frac{(x_1-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2} \right)}{2(1-\rho^2)} \right]$$

The joint PDF can be written in matrix form as

$$f_{X_1X_2}(x_1, x_2) = \frac{1}{2\pi\sqrt{|\det(\Sigma)|}} \exp\left[\underbrace{-\frac{1}{2}\underline{(x-\mu)}^T\Sigma^{-1}\underline{(x-\mu)}} \right]$$

where

covariance matrix

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$$

- Notation for bivariate normal: $\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \sim \mathcal{N}(\mu, \Sigma)$

Linear Transformation of 2 Random Variables

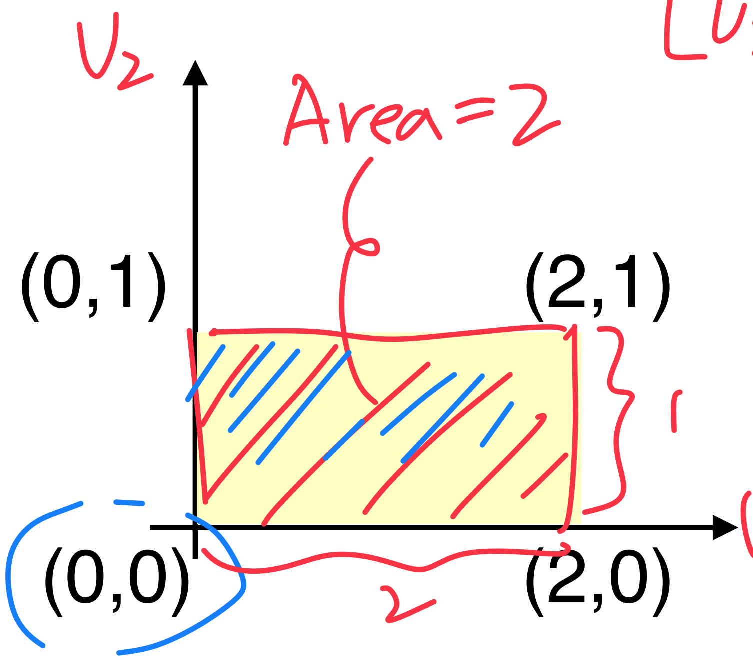
- **Theorem**: Let U_1, U_2, V_1, V_2 be random variables that satisfy $V_1 = aU_1 + bU_2$ and $V_2 = cU_1 + dU_2$. Define the matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Assume that $\det(A) \neq 0$. Then, we have

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} U_1 \\ U_2 \end{bmatrix}$$

$$f_{V_1V_2}(v_1, v_2) = \frac{1}{|\det(A)|} f_{U_1U_2}(A^{-1}[v_1, v_2]^T)$$

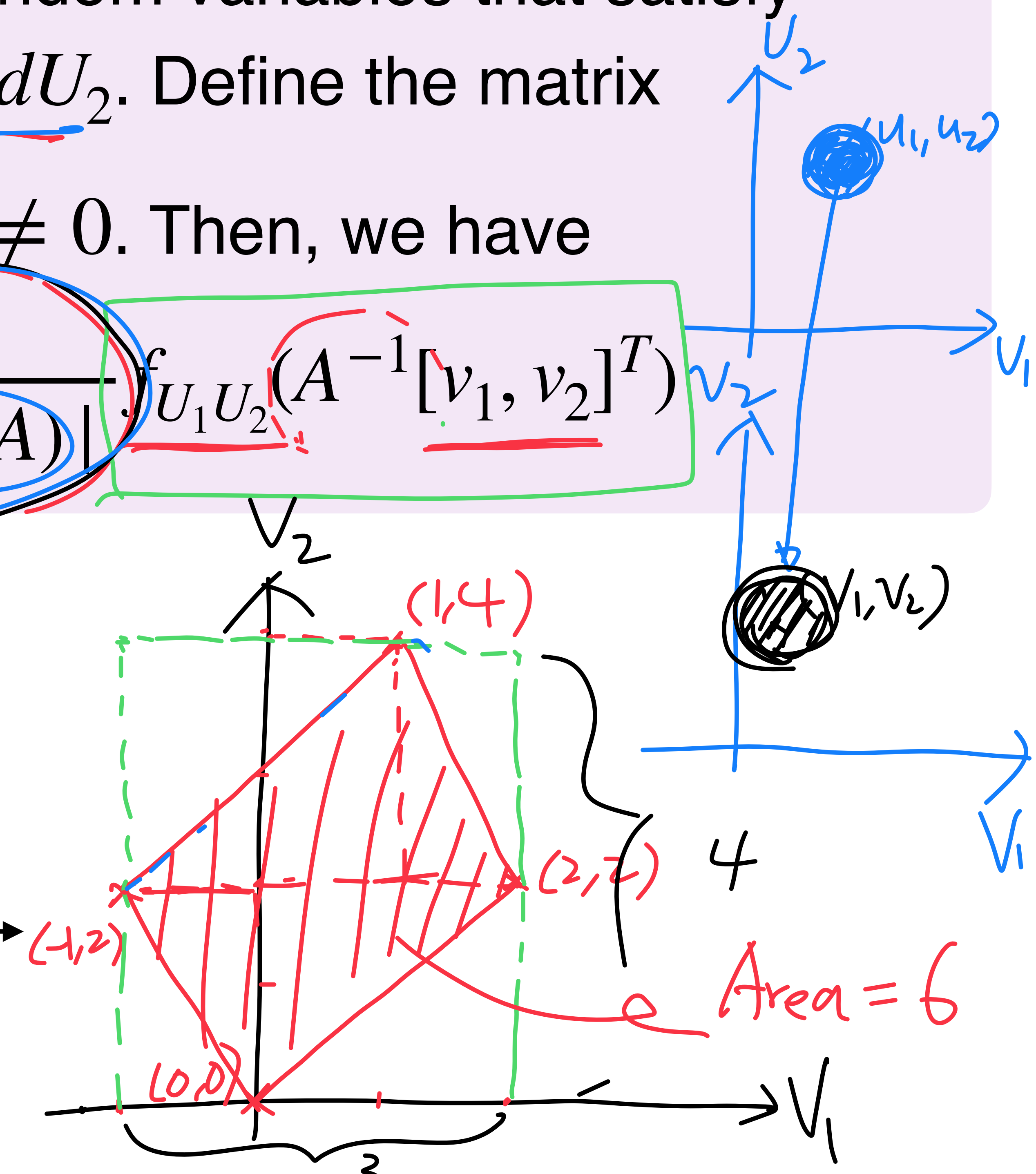
$$\rightarrow \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} = A^{-1} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix}$$

- **Intuition**:



$$A = \begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix}$$

det(A)=3



For more details, please check: **https://www.stat.berkeley.edu/~aditya/resources/AllLectures2018Fall201A.pdf**

Bivariate Normal and Linear Transformation

Z, W standard normal

- For simplicity, assume $\mu_1 = \mu_2 = 0$ (can be handled via translation)

$$\begin{cases} X_1 = \sigma_1 Z \\ X_2 = \sigma_2\left(\rho Z + \sqrt{1-\rho^2}\,W\right) \end{cases}$$

$$f_{X_1X_2}(x_1, x_2) = \frac{1}{|\det(A)|} f_{ZW}(A^{-1}[x_1, x_2]^T)$$

$$\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \underbrace{\begin{bmatrix} \sigma_1 & 0 \\ \sigma_2 \cdot \rho & \sigma_2 \cdot \sqrt{1-\rho^2} \end{bmatrix}}_{\shortparallel \atop A} \begin{bmatrix} Z \\ W \end{bmatrix}$$

$$\det(A) = \sigma_1 \sigma_2 \cdot \sqrt{1-\rho^2}$$

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} \sigma_2\sqrt{1-\rho^2} & 0 \\ -\sigma_2 \cdot \rho & \sigma_1 \end{bmatrix}$$

$$f_{ZW}(z, w) = f_Z(z) \cdot f_W(w)$$

$$= \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{z^2}{2\sigma_1^2}} \cdot \frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-\frac{w^2}{2\sigma_2^2}}$$

$$X_2 = \sigma_2\left(\rho Z + \sqrt{1-\rho^2}\, W\right) + \mu_2 \quad , \quad X_2 \sim N(\mu_2, \sigma_2^2)$$

There are still a few remaining questions:

(Q1) Is X_2 a normal random variable? What is the PDF?

Sum of independent random variables

(Q2) What is "ρ" in the joint PDF of bivariate normal?

Covariance

X_1, X_2 $\quad X_1$ $\quad X_2$

(Q3) Why is bivariate normal useful? Any nice properties?

Conditional PDF and beyond

$f_{X_1|X_2}$ normal

$f_{X_2|X_1}$

(Q1) Sum of Independent Random Variables and Moment Generating Functions (MGF)

$Z = X + Y$ and X, Y Independent — Discrete Case

- Question: X, Y are two independent discrete random variables.
  - Define $Z = X + Y$
  - What's the PMF of Z?

卷積

Convolution Theorem: Let X, Y be two independent discrete random variables with PMF $p_X(x)$ and $p_Y(y)$. Define $Z = X + Y$. Then, the PMF of Z is

$$p_Z(z) = P(Z = z) = \sum_x p_X(x) p_Y(z - x)$$

- Recall: $X \sim \text{Poisson}(\lambda_1, T)$, $Y \sim \text{Poisson}(\lambda_2, T)$, $Z = X + Y$
  - What's the PMF of Z?

Z

$Z = X + Y$ and X, Y Independent — Continuous Case

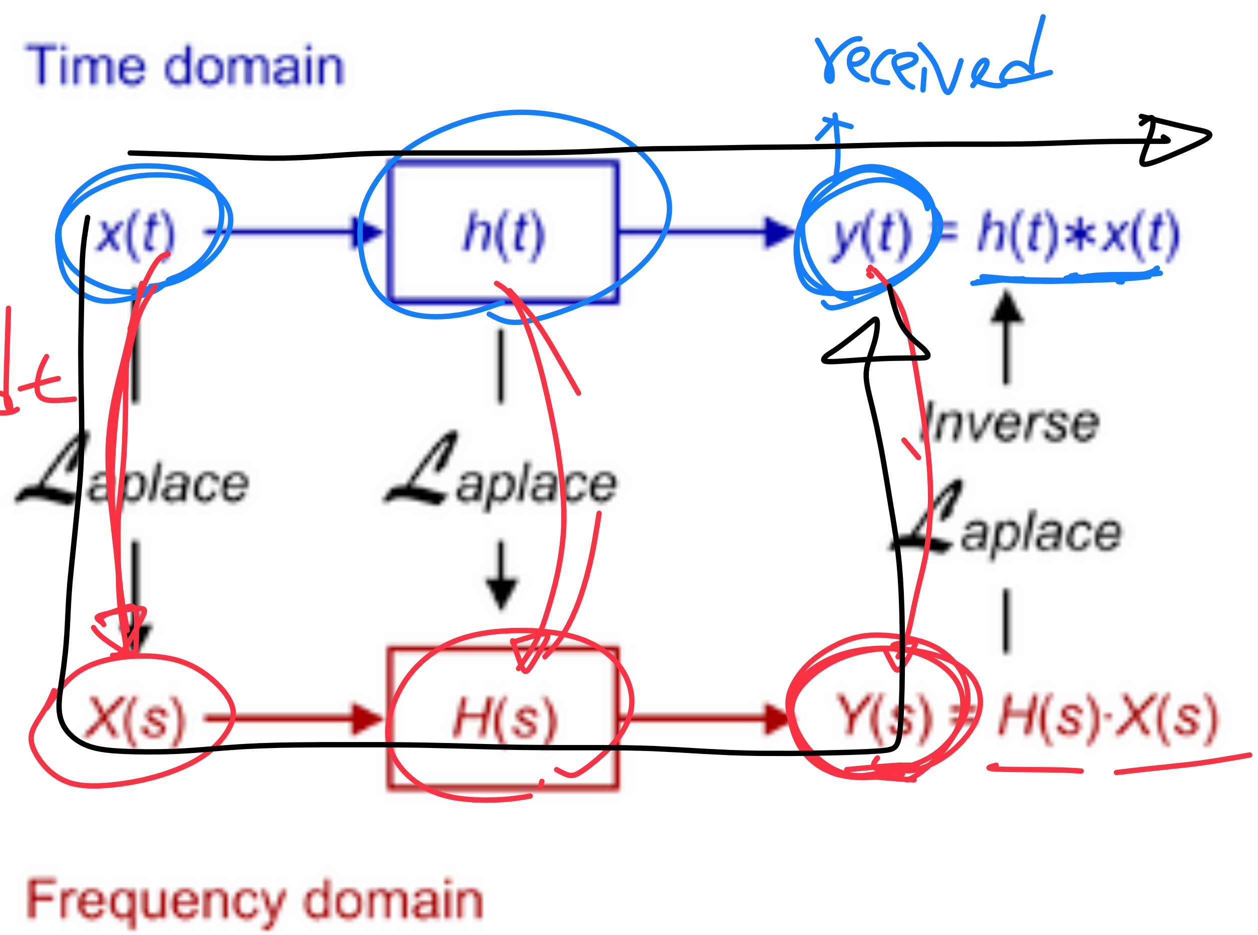
- For continuous random variables:

Convolution Theorem: Let X, Y be two continuous independent random variables with PDF f_1 and f_2. Define $Z = X + Y$. Then, the PDF of Z is

$$f_Z(z) = \int_{-\infty}^{\infty} f_1(x) f_2(z - x) dx$$

Any Issue With Convolution Theorem?

- Issue: Sometimes it is quite tedious to do convolution
- Question: Any other approach?
- Idea: Borrow ideas from signal processing — "Laplace transform"



$$X(s) = \int_{-\infty}^{+\infty} x(t)\cdot e^{-st}\,dt$$

- In Probability, this is called "Moment Generating Function"

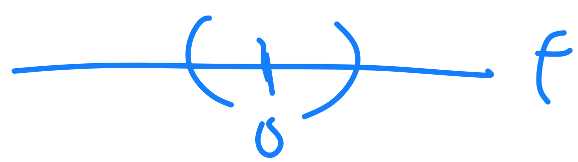
Moment Generating Function (Formally)

- **Moment Generating Function (MGF)**: For a random variable X, define

$$M_X(t) = E[e^{tX}], \quad t \in \mathbb{R}$$

If there exists $\delta > 0$ such that $M_X(t) < \infty$ for all $t \in (-\delta, \delta)$, then $M_X(t)$ is called the moment generating function of X

- Remark: If X is discrete with PMF $p_X(x)$, then



$$M_X(t) = \sum_{\text{all } x} e^{tx} \cdot p_X(x)$$

- Remark: If X is continuous with PDF $f_X(x)$, then

$$M_X(t) = \int_{-\infty}^{+\infty} f_X(x)\, e^{tx}\, dx$$

Example: Find MGF of Normal Random Variables

- Example: Let $Z \sim \mathcal{N}(\mu, \sigma^2)$
  - Question: What is the MGF of Z, $M_Z(t)$ = ?

$$M_Z(t) = E\left[e^{tZ}\right] = \int_{-\infty}^{+\infty} e^{tz} \cdot \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(z-\mu)^2}{2\sigma^2}} \, dz$$

$$= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left(-\frac{(1)z^2 + (-2\mu + 2t\sigma^2)z + (\mu^2)}{2\sigma^2}\right) dz$$

PDF of $\mathcal{N}(\mu+t\sigma^2, \sigma^2)$

$$= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left(-\frac{(z-(\mu+t\sigma^2))^2}{2\sigma^2} + \frac{(\mu+t\sigma^2)^2 - \mu^2}{2\sigma^2}\right) dz$$

does not involve z

$$= \exp\left(\frac{1}{2}t^2\sigma^2 + \mu t\right)$$

Nice Properties of MGF?

- **Let X_1, X_2 be two random variables**:

1. ______ Suppose $M_{X_1}(t) = M_{X_2}(t)$, for all $t \in \mathbb{R}$. Do X_1 and X_2 always have the same distribution (i.e., the same CDF)?

2. ______ Could we find moments $E[X_1^n]$ by using $M_{X_1}(t)$?

3. ______ Suppose X_1, X_2 are independent. Could we express $M_{X_1+X_2}(t)$ in terms of $M_{X_1}(t), M_{X_2}(t)$?

Nice Property (I): MGF Uniqueness Theorem

- **MGF Uniqueness Theorem**: Let X_1 and X_2 be two random variables with MGFs $M_{X_1}(t)$ and $M_{X_2}(t)$, respectively. If $M_{X_1}(t) = M_{X_2}(t)$ for all t in some interval $(-\alpha, \alpha)$, then X_1 and X_2 follow the same distribution, i.e.

$$P(X_1 \leq u) = P(X_2 \leq u), \text{ for all } u \in \mathbb{R}$$

- Remark: More details in the following reference
  - J. H. Curtiss, "A note on the theory of moment generating functions," 1942
  - https://projecteuclid.org/download/pdf_1/euclid.aoms/1177731541

Example: Find CDF from MGF

- Example: Suppose the MGF of a random variable X is
  $$M_X(t) = \frac{1}{6}e^{-2t} + \frac{1}{3}e^{-t} + \frac{1}{4}e^{t} + \frac{1}{4}e^{2t}$$
  - Question: $P(|X| \leq 1)$ = ?

MGF of Special Random Variables

Distribution	Moment-generating function $M_X(t)$
Degenerate δ_a	e^{ta}
Bernoulli $P(X=1)=p$	$1-p+pe^t$
Geometric $(1-p)^{k-1}\,p$	$\dfrac{pe^t}{1-(1-p)e^t}$ $\forall t < -\ln(1-p)$
Binomial $B(n,p)$	$(1-p+pe^t)^n$
Negative Binomial $NB(r,p)$	$\dfrac{(1-p)^r}{(1-pe^t)^r}$
Poisson $Pois(\lambda)$	$e^{\lambda(e^t-1)}$
Uniform (continuous) $U(a,b)$	$\dfrac{e^{tb}-e^{ta}}{t(b-a)}$
Uniform (discrete) $DU(a,b)$	$\dfrac{e^{at}-e^{(b+1)t}}{(b-a+1)(1-e^t)}$
Laplace $L(\mu,b)$	$\dfrac{e^{t\mu}}{1-b^2t^2},\ \lvert t\rvert < 1/b$
Normal $N(\mu,\sigma^2)$	$e^{t\mu+\frac{1}{2}\sigma^2t^2}$

- Example: If $M_X(t) = \dfrac{1}{2} + \dfrac{1}{2}e^t$, then what kind of r.v. is X?

- Example: If $M_Z(t) = e^{2t^2-t}$, then what kind of r.v. is Z?

Nice Property (II): From Sum to Product

- **MGF and Sum of 2 Independent Random Variables**: Given 2 independent random variables X_1, X_2 with MGFs $M_{X_1}(t)$ and $M_{X_2}(t)$, the MGF of $X_1 + X_2$ is

$$M_{X_1+X_2}(t) = M_{X_1}(t) \cdot M_{X_2}(t)$$

- Proof:

Example: MGF of Sum of 2 Normal R.V.s

- Example: $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$ and $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$
  - X_1 and X_2 are assumed to be independent
  - Question: What is the MGF of $X_1 + X_2$? What is the PDF of $X_1 + X_2$?

Nice Property (III): Why Is $M_X(t)$ Called the Moment Generating Function?

- Recall: What is the "n-th moment" of X?

- **Use MGF to Find Moments**: Let X be a random variable with MGF $M_X(t)$. Then, for every $n \in \mathbb{N}$, we have

$$E[X^n] = \frac{d^n}{dt^n} M_X(t)\big|_{t=0}$$

- Proof:

Example: Moments of Exp(λ)

- Example: Suppose $X \sim$ Exp(λ)
  - What is the MGF of X?
  - Use MGF to verify that $E[X] = \dfrac{1}{\lambda}$ and Var$[X] = \dfrac{1}{\lambda^2}$?