Lecture 9

- **Read:** Chapter 4.11, 6.1-6.6.
- Multiple Continuous Random Variables
  - Jointly (Bivariate) Gaussian Random Variables
- Sums of Random Variables
  - Expectations of Sums
  - PDF of the Sum of Two Random Variables
  - Moment Generating Function
  - Moment Generating Function of the Sum of Independent Random Variables
  - Sums of Independent Gaussian Random Variables
  - Sum of a Random Number of Independent Random Variables

Jointly Gaussian Random Variables

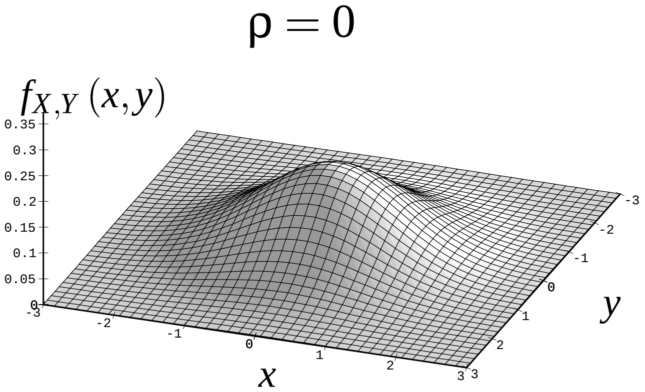
- **Definition:** X and Y have a bivariate Gaussian PDF if

$$f_{X,Y}(x,y) = \frac{exp\left[-\frac{\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - \frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \left(\frac{y-\mu_2}{\sigma_2}\right)^2}{2(1-\rho^2)}\right]}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}$$

where $\sigma_1 > 0$, $\sigma_2 > 0$, $-1 \leq \rho \leq 1$

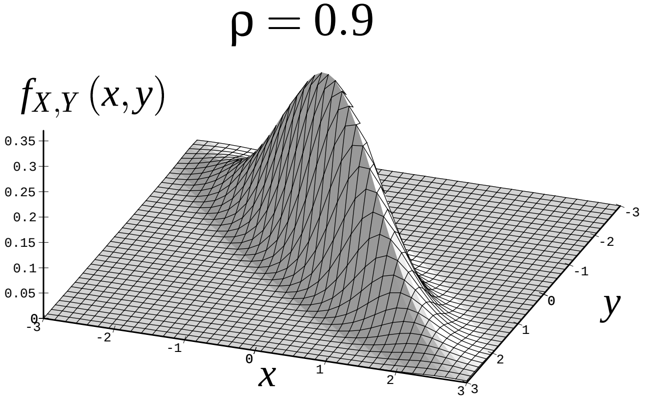
When $\rho = 0$

- $\mu_1 = \mu_2 = 0$, $\sigma_1 = \sigma_2 = 1$
- Joint PDF has circular symmetry of a hat



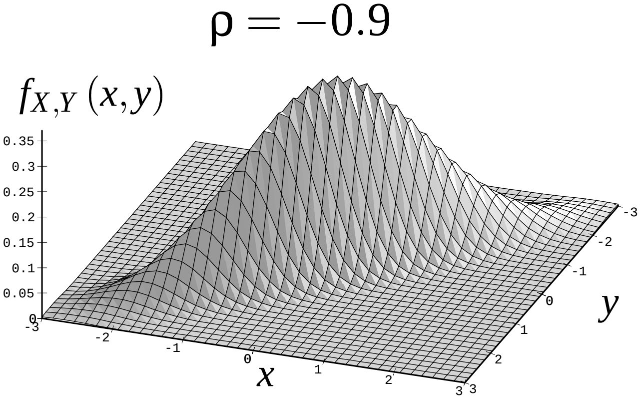
When $\rho = 0.9$

- $\mu_1 = \mu_2 = 0$, $\sigma_1 = \sigma_2 = 1$
- Joint PDF forms a ridge over the line $x = y$
- The ridge becomes increasingly steep as $\rho \to 1$



When $\rho = -0.9$

- $\mu_1 = \mu_2 = 0$, $\sigma_1 = \sigma_2 = 1$
- Joint PDF forms a ridge over the line $x = -y$
- The ridge becomes increasingly steep as $\rho \to -1$



Rewriting the Bivariate Gaussian PDF

- Complete the square of the exponent to write

$$f_{X,Y}(x,y) = f_X(x) f_{Y|X}(y|x)$$

where

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-(x-\mu_1)^2/2\sigma_1^2}$$
$$f_{Y|X}(y|x) = \frac{1}{\sqrt{2\pi}\tilde{\sigma}_2} e^{-(y-\tilde{\mu}_2(x))^2/2\tilde{\sigma}_2^2}$$

Bivariate Gaussian Properties

- $E[X] = \mu_1$
- Given $X = x$, Y is Gaussian
- Conditional mean of Y given $X = x$:

$$\begin{aligned} \tilde{\mu}_2(x) &= \mu_2 + \rho \frac{\sigma_2}{\sigma_1}(x - \mu_1) \\ &= E[Y|X = x] \end{aligned}$$

Gaussian Marginal PDF When $\rho = 0$ (X and Y are Uncorrelated)

- **Theorem:** If X and Y are the bivariate Gaussian random variables in our definition above and $\rho = 0$

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-(x-\mu_1)^2/2\sigma_1^2}$$
$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_2} e^{-(y-\mu_2)^2/2\sigma_2^2}$$

Gaussian Conditional PDF

- Given the marginal PDFs of X and Y, we use the definition of the conditional PDF to find the conditional PDFs.
- If X and Y are the bivariate Gaussian random variables defined above, the conditional PDF of Y given X is

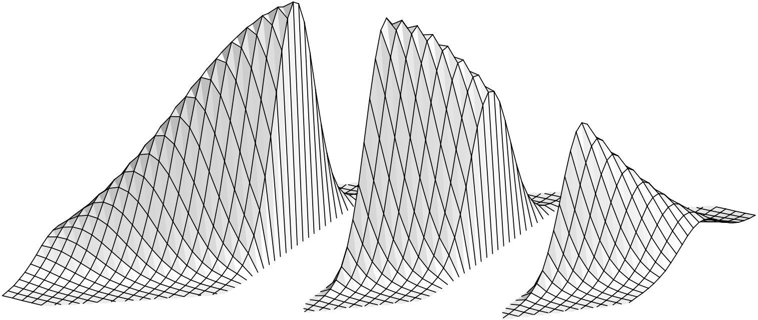
$$f_{Y|X}(y|x) = \frac{1}{\sqrt{2\pi}\tilde{\sigma}_2} e^{-(y-\tilde{\mu}_2(x))^2/2\tilde{\sigma}_2^2}$$

where

$$\tilde{\mu}_2(x) = E[Y|X=x] = \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)$$
$$\tilde{\sigma}_2^2 = Var[Y|X=x] = \sigma_2^2(1-\rho^2)$$

Gaussian Conditional PDF

- Cross-sectional view of the joint Gaussian PDF with $\mu_1 = \mu_2 = 0$, $\sigma_1 = \sigma_2 = 1$, and $\rho = 0.9$
- The bell shape of the cross section occurs because the conditional PDF $f_{Y|X}(y|x)$ is Gaussian.



More Than Two Continuous RVs

- **Definition: (Multivariate Joint CDF)** The joint CDF of $X_1, ..., X_n$ is

$$F_{X_1,...,X_n}(x_1, ..., x_n) = P[X_1 \leq x_1, ..., X_n \leq x_n]$$

- **Definition: (Multivariate Joint PDF)** The joint PDF of $X_1, ..., X_n$ is $f_{X_1,...,X_n}(x_1, ..., x_n)$ satisfying

$$F_{X_1,...,X_n}(x_1, ..., x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f_{X_1,...,X_n}(u_1, ..., u_n) du_1 ... du_n$$

Joint PDF Properties

- $f_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = \frac{\partial^n F_{X_1,\ldots,X_n}(x_1,\ldots,x_n)}{\partial x_1 \ldots \partial x_n}$
- $f_{X_1,\ldots,X_n}(x_1,\ldots,x_n) \geq 0$
- $\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{X_1,\ldots,X_n}(x_1,\ldots,x_n) dx_1 \ldots dx_n = 1$
- $P[A] = \underset{A}{\int \cdots \int} f_{X_1,\ldots,X_n}(x_1,\ldots,x_n) dx_1 dx_2 \ldots dx_n$

Marginal PDFs

- **Theorem:** For a joint PDF of four random variables, $f_{W,X,Y,Z}(w,x,y,z)$, some marginal PDFs are

$$
\begin{aligned}
f_{X,Y,Z}(x,y,z) &= \int_{-\infty}^{\infty} f_{W,X,Y,Z}(w,x,y,z)dw \\
f_{W,Z}(w,z) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{W,X,Y,Z}(w,x,y,z)dxdy \\
f_X(x) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{W,X,Y,Z}(w,x,y,z)dwdydz
\end{aligned}
$$

- Can be generalized in a straightforward way to any marginal PDF of a joint PDF of an arbitrary number of random variables.

N Independent Random Variables

- **Definition: (N Independent Random Variables)** $X_1, ..., X_n$ are independent if

$$f_{X_1,...,X_n}(x_1, ..., x_n) = f_{X_1}(x_1)...f_{X_n}(x_n)$$

for all $x_1, ..., x_n$.

N Independent Random Variables

- Mutual independence of n random variables is typically the result of an experiment with special structure that ensures the independence
- The most common example occurs when an experiment consists of n independent trials.
- In this case, trial i produces the random variable X_i. Since all trials follow the same experiment, all of the X_i have the same PDF. In this case, we say the random variables X_i are identically distributed.
- **Definition: (Independent and Identically Distributed)** $X_1, ..., X_n$ are independent and identically distributed (iid) if and only if

$$f_{X_1,...,X_n}(x_1, ..., x_n) = f_X(x_1)...f_X(x_n)$$

for all $x_1, ..., x_n$.

Function of N Random Variables

- Just as we did for one random variable and two random variables, we can derive a new random variable $Y = g(X_1, ..., X_n)$ that is a function of n random variables.
- When the X_i are continuous, we can find the CDF of Y

$$F_Y(y) = P[Y \leq y] = \int \cdots \int_{g(x_1,...,x_n) \leq y} f_{X_1,...,X_n}(x_1, ..., x_n) dx_1 ... dx_n$$

Expectation of a Function of N Random Variables

- **Theorem:** For $Y = g(X_1, ..., X_n)$, the expected value is

$$\begin{aligned} E[Y] &= E[g(X_1, ..., X_n)] \\ &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(x_1, ..., x_n) f_{X_1,...,X_n}(x_1, ..., x_n) dx_1 ... dx_n \end{aligned}$$

- When $(X_1, ..., X_n)$ are independent, the expected value of $g(X_1) \times \cdots \times g(X_n)$ is the product of the expected values.

- **Theorem:** If $X_1, ..., X_n$ are independent random variables,

$$E[g(X_1)g(X_2) \cdots g(X_n)] = E[g(X_1)]E[g(X_2)] \cdots E[g(X_n)]$$

N Random Variables: Example 1

- Let $X_1, ..., X_n$ be iid RVs, with mean 0, variance 1 and covariance $Cov[X_i, X_j] = \rho$.
- Find the expected value and variance of the sum $Y = X_1 + ... + X_n$.

....

- The mean value of a sum of random variables is always the sum of their individual means.

$$E[Y] = \sum_{i=1}^{n} E[X_i] = 0$$

N Random Variables: Example 1 (cont.)

- The variance of any sum of random variables can be expressed in terms of the individual variances and covariances.

- Since $E[Y]$ is zero, $Var[Y] = E[Y^2]$. Thus,

$$Var[Y] = E\left[\left(\sum_{i=1}^{n} X_i\right)^2\right] = E\left[\sum_{i=1}^{n}\sum_{j=1}^{n} X_i X_j\right]$$

$$= \sum_{i=1}^{n} E[X_i^2] + \sum_{i=1}^{n}\sum_{j\neq i}^{n} E[X_i X_j]$$

- Since $E[X_i] = 0$, $E[X_i^2] = Var[X_i] = 1$ and for $i \neq j$

$$E[X_i X_j] = Cov[X_i, X_j] = \rho$$

- Thus,

$$Var[Y] = n + n(n-1)\rho$$

N Random Variables: Example 2

- Let $X_1, ..., X_n$ denote n iid random variables each with PDF $f_X(x)$.
- Find the CDF and PDF of $Y = min(X_1, ..., X_n)$.

. . . .

N Random Variables: Example 2 (cont.)

- We have

$$\begin{aligned} P[Y \geq y] &= P[min(X_1, ..., X_n) \geq y] \\ &= P[X_1 \geq y, ..., X_n \geq y] \\ &= (P[X_1 \geq y])^n \\ &= [1 - F_X(y)]^n \end{aligned}$$

- Therefore, the CDF is

$$\begin{aligned} F_Y(y) = P[Y \leq y] &= 1 - P[Y \geq y] \\ &= 1 - (1 - F_X(y))^n \end{aligned}$$

- So, the PDF is

$$f_Y(y) = \frac{dF_Y(y)}{dy} = n(1 - F_X(y))^{n-1} f_X(y)$$

Sums of Random Variables

- Wide variety of questions can be answered by studying a random variable, W_n defined as sum of n random variables:

$$W_n = X_1 + ... + X_n$$

- Since W_n is a function of n random variables, we could refer to the joint distribution of $X_1, ..., X_n$ to derive the complete probability model of W_n in the form of a PMF or PDF.

Sums of Random Variables (cont.)

- However, in many practical applications, the nature of the analysis or the properties of the random variables allow us to apply techniques that are simpler than analyzing a general n-dimensional probability model
  - for $E[W_n]$ and $Var[W_n]$
  - when $X_1, ..., X_n$ iid (independent and identically distributed)
- These techniques for sums of independent random variables
  - will allow us to calculate moments and derive relationships between families of random variables
  - apply to both discrete and continuous random variables

Expectations of Sums

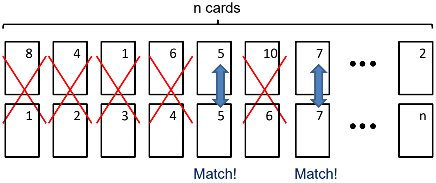
- **Theorem:** For **any** set of random variables $X_1, ..., X_n$, the expected value of $W_n = X_1 + ... + X_n$ is

$$E[W_n] = E[X_1] + E[X_2] + ... + E[X_n]$$

- **Proof:** By induction on n
- **Note:** The expectation of the sum equals the sum of the expectations whether or not $X_1, ..., X_n$ are independent!

Example: Matching Cards

- Label a deck of n cards $1, \ldots, n$.
- Shuffle and turn over one at a time.



- $X_i = 1$ if the ith card is labeled i.
- Number of matches is

$$W = X_1 + \ldots + X_n$$

- Find $E[W]$.

. . . .

- Since the probability of a card matching its label is $1/n$,

$$P[X_i = 1] = 1/n$$

- So, $E[X_i] = 1/n$.

$$\begin{aligned} E[W] &= E[X_1] + \ldots + E[X_n] \\ &= nE[X_i] = n \cdot 1/n = 1 \end{aligned}$$

Example: Matching Cards (cont.)

- **Note:** It is tempting to think that $P[X_i = 1]$ should change as we turn over more cards.
- That is, only the first card will have a $1/n$ probability of matching its label.
- The second card would then have $1/(n-1)$, and so forth.
- This line of reasoning is wrong!

Example: Matching Cards (cont.)

- The second card would have $1/(n-1)$ probability, given the fact that its label did not come up on the first card.
- If the first card revealed the label 2, then the second card has a probability of 0.
- Consequently, when all possible outcomes are considered, the probability is always $1/n$ for each card.

$$
\begin{aligned}
P[X_2 = 1] &= P[X_2 = 1 | \text{1st} \neq \text{label } 2] P[\text{1st} \neq \text{label } 2] \\
&\quad + P[X_2 = 1 | \text{1st} = \text{label } 2] P[\text{1st} = \text{label } 2] \\
&= \frac{1}{n-1} \cdot \frac{n-1}{n} + 0 \cdot \frac{1}{n} \\
&= \frac{1}{n}
\end{aligned}
$$

Variance of Sums

- **Theorem:** The variance of $W_n = X_1 + ... + X_n$ is

$$Var[W_n] = \sum_{i=1}^{n} Var[X_i] + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} Cov[X_i, X_j]$$

  (Proof by algebra)

- **Theorem:** When $X_1, ..., X_n$ are mutually independent, the variance of $W_n = X_1 + ... + X_n$ is the sum of the variances:

$$Var[W_n] = Var[X_1] + ... + Var[X_n]$$

  (Because when $X_1, ..., X_n$ are mutually independent, the terms $Cov[X_i, X_j] = 0$ if $i \neq j$.)

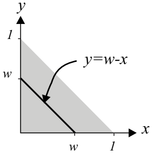
Variance of Sums: Example

- X and Y have the joint PDF

$$f_{X,Y}(x,y) = \begin{cases} 2 & , 0 \leq y \leq 1,\ 0 \leq x \leq 1,\ x+y \leq 1 \\ 0 & , \text{otherwise} \end{cases}$$

- Find the variance of $W = X + Y$.

. . . .



- According to the theorem:

$$Var[W] = Var[X] + Var[Y] + 2Cov[X,Y]$$

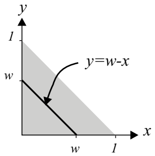
- First two moments of X:

$$E[X] = \int_0^1 \int_0^{1-x} 2x\,dy\,dx = \int_0^1 2x(1-x)dx = 1/3$$

$$E[X^2] = \int_0^1 \int_0^{1-x} 2x^2\,dy\,dx = \int_0^1 2x^2(1-x)dx = 1/6$$

- So, X has variance, $Var[X] = E[X^2] - (E[X])^2 = 1/18$. By symmetry, $E[Y] = E[X] = 1/3$, $Var[Y] = Var[X] = 1/18$.

Variance of Sums: Example (cont.)



- To find the covariance, we first find the correlation:

$$E[XY] = \int_0^1 \int_0^{1-x} 2xydydx = \int_0^1 x(1-x)^2 dx = 1/12$$

- Covariance is:

$$Cov[X,Y] = E[XY] - E[X]E[Y] = 1/12 - (1/3)^2 = -1/36$$

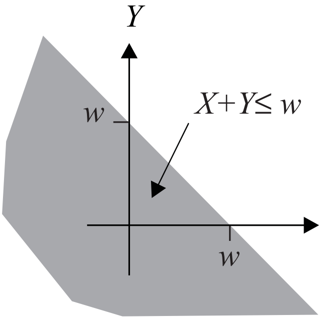
- Finally, the variance of the sum $W = X + Y$ is

$$\begin{aligned} Var[W] &= Var[X] + Var[Y] + 2Cov[X,Y] \\ &= 2/18 - 2/36 = 1/18 \end{aligned}$$

PDF of the Sum of Two Random Variables

- **Theorem:** The PDF of $W = X + Y$ is

$$f_W(w) = \int_{-\infty}^{\infty} f_{X,Y}(x, w-x)dx$$
$$= \int_{-\infty}^{\infty} f_{X,Y}(w-y, y)dy$$



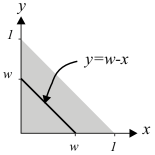
PDF of the Sum of Two Random Variables: Example

- Find the PDF of $W = X + Y$ when X and Y have the joint PDF

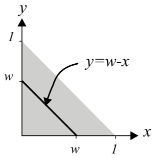
$$f_{X,Y}(x,y) = \begin{cases} 2 & , 0 \leq y \leq 1,\ 0 \leq x \leq 1,\ x + y \leq 1 \\ 0 & , \text{otherwise} \end{cases}$$

. . . .

- PDF of $W = X + Y$ can be found using the theorem
- X and Y are dependent, and possible values of X, Y occur in the shaded triangular region ($0 \leq X + Y \leq 1$).



PDF of the Sum of Two Random Variables: Example (cont.)



- Thus, $f_W(w) = 0$ for $w < 0$ or $w > 1$.
- For $0 \leq w \leq 1$, applying the theorem yields

$$f_W(w) = \int_0^w 2dx = 2w \qquad (0 \leq w \leq 1)$$

- The complete expression for the PDF of W is

$$f_W(w) = \begin{cases} 2w & , 0 \leq w \leq 1 \\ 0 & , \text{otherwise} \end{cases}$$

PDF of the Sum of Two Independent Random Variables

- **Theorem:** When X and Y are independent random variables, the PDF of $W = X + Y$ is

$$f_W(w) = \int_{-\infty}^{\infty} f_X(w-y) f_Y(y)\,dy$$
$$= \int_{-\infty}^{\infty} f_X(x) f_Y(w-x)\,dx$$

- PDF of an independent sum is the convolution of the PDFs.

PDF of the Sum of Two Independent Random Variables

- Convolution notation: $f_W(w) = f_X(x) * f_Y(y)$
- It is often helpful to use transform methods to compute the convolution of two functions.
- In the language of probability theory, the transform of a PDF or a PMF is a moment generating function (MGF).
- Convolution of the PDFs is equivalent to multiplication of MGFs
- Summing RVs is equivalent to multiplying MGFs

Related Courses

- MAT 210E Engineering Mathematics (differential equations, Laplace transform) (3rd semester)
- BLG 354E Signal and Systems for Computer Engineering (Laplace transform, z-transform, Fourier transform, convolution) (6th semester)
- EHB 252E Signals and Systems (z-transform, Fourier transform, convolution) (4th semester)

Moment Generating Function (MGF)

- **Definition: (Moment Generating Function)** For a random variable X, the moment generating function (MGF) of X is

$$\phi_X(s) = E\left[e^{sX}\right]$$

- If X is a continuous random variable

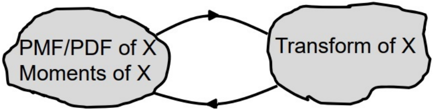
$$\phi_X(s) = \int_{-\infty}^{\infty} e^{sx} f_X(x)dx$$

  - This equation indicates that the MGF of a continuous random variable is similar to the Laplace transform of a time function.

- If Y is a discrete random variable

$$\phi_Y(s) = \sum_{y_i \in S_Y} e^{sy_i} p_Y(y_i)$$

Why Use Transforms?



- A different way of representing distribution of RV
- Ease of computation in transformed space
  - Calculation of moments
  - Distributions of random sums of RVs
  - Analytical derivations and theorem proving

Transforms

- When do they exist?
- Properties
- Inversion, i.e., back to PMF/PDF

Region of Convergence (I)

- In the integral form, the MGF is reminiscent of the Fourier and Laplace transforms that are commonly used in linear systems.

$$\phi_X(s) = \int_{-\infty}^{\infty} e^{sx} f_X(x) dx$$

- The primary difference is that the MGF is defined for real values of s.
- For a given random variable X, there is a range of possible values of s for which $\phi_X(s)$ exists.
- The set of values of s for which $\phi_X(s)$ exists is called the region of convergence.

Region of Convergence (II)

- For example, if X is a nonnegative random variable, the region of convergence includes all $s \leq 0$.
- For any random variable X, $\phi_X(s)$ always exists for $s = 0$.
- We will use the moment generating function by evaluating its derivatives at $s = 0$.
- As long as the region of convergence includes a nonempty interval $(-\epsilon, \epsilon)$ about the origin $s = 0$, we can evaluate the derivatives of the MGF at $s = 0$.
- This is the case for commonly used random variables.

MGF as a Complete Model

- Like the PMF of a discrete random variable and the PDF of a continuous random variable, the MGF is a complete probability model of a random variable.
- Using inverse transform methods, it is possible to calculate the PMF or PDF from the MGF.

MGF ($\phi_X(s)$) Properties

- **Theorem:** For any random variable X, the MGF satisfies
$$\phi_X(s)|_{s=0} = 1$$
  - The definition of the MGF implies this: $\phi_X(0) = E[e^{0 \cdot X}] = E[1] = 1$.
  - This theorem is quite useful in checking that an alleged MGF $\phi_X(s)$ is valid.
- **Theorem: (Linear Function of an RV)** The MGF of $Y = aX + b$ satisfies
$$\phi_Y(s) = e^{sb}\phi_X(as)$$
- As its name suggests, the function $\phi_X(s)$ is especially useful for finding the moments of X.
- **Theorem: (From Transforms to Moments)** A random variable X with MGF $\phi_X(s)$ has nth moment
$$E[X^n] = \left.\frac{d^n\phi_X(s)}{ds^n}\right|_{s=0}$$

Moment Generating Functions of Families of Random Variables

- In Appendix A of our textbook, under the definition of each random variable, the MGF of that random variable is given.

MGF Examples

- **Example 1:** If $X = a$ (a constant), then $f_X(x) = \delta(x - a)$ and

$$\phi_X(s) = \int_{-\infty}^{\infty} e^{sx}\delta(x - a)dx = e^{sa}$$

- **Example 2:** When X has the uniform PDF

$$f_X(x) = \begin{cases} 1 & , 0 \leq x \leq 1 \\ 0 & , \text{otherwise} \end{cases}$$

the moment generating function of X is

$$\phi_X(s) = \int_0^1 e^{sx}dx = \frac{e^s - 1}{s}$$

MGF of Exponential RV

- Let X have the exponential PDF

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & , x \geq 0 \\ 0 & , \text{otherwise} \end{cases}$$

The MGF of X is

$$\begin{aligned} \phi_X(s) &= \int_0^\infty e^{sx} \lambda e^{-\lambda x} dx = \lambda \int_0^\infty e^{(s-\lambda)x} dx \\ &= \lambda \frac{e^{(s-\lambda)x}}{s-\lambda} \Big|_0^\infty \\ &= \frac{\lambda}{\lambda - s}, \text{ if } s < \lambda \end{aligned}$$

MGF of Bernoulli RV

- Let X be a Bernoulli random variable with

$$p_X(x) = \begin{cases} 1-p & , x = 0 \\ p & , x = 1 \\ 0 & , \text{otherwise} \end{cases}$$

The MGF of X is

$$\phi_X(s) = E\left[e^{sX}\right] = (1-p)e^0 + pe^s = 1 - p + pe^s$$

MGF of Geometric RV

- Let N have a geometric PMF

$$p_N(n) = \begin{cases} (1-p)^{n-1}p & , n = 1, 2, \ldots \\ 0 & , \text{otherwise} \end{cases}$$

The MGF of N is

$$\begin{aligned} \phi_N(s) &= \sum_{n=1}^{\infty} e^{sn} p(1-p)^{n-1} \\ &= pe^s \sum_{n=1}^{\infty} [(1-p)e^s]^{n-1} \quad \text{(sum of an infinite geometric series)} \\ &= \frac{pe^s}{1-(1-p)e^s} \end{aligned}$$

MGF of Poisson RV

- Let K have the Poisson PMF

$$p_K(k) = \begin{cases} \frac{\alpha^k e^{-\alpha}}{k!} & , k = 0, 1, \ldots \\ 0 & , \text{otherwise} \end{cases}$$

The MGF of K is

$$\begin{aligned} \phi_K(s) &= \sum_{k=0}^{\infty} e^{sk} \alpha^k e^{-\alpha} / k! \\ &= e^{-\alpha} \sum_{k=0}^{\infty} (\alpha e^s)^k / k! \quad \text{(power series expansion of exp. function)} \\ &= e^{\alpha(e^s - 1)} \end{aligned}$$

MGF of Gaussian RV (I)

- **Theorem:** If $Z \sim N(0,1)$, then the MGF of Z is

$$\phi_Z(s) = e^{s^2/2}$$

- **Proof:** MGF of Z is

$$\begin{aligned}\phi_Z(s) &= \int_{-\infty}^{\infty} e^{sz} f_Z(z) dz \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{sz} e^{-z^2/2} dz\end{aligned}$$

- Completing the square in the exponent:

$$\begin{aligned}\phi_Z(s) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(z^2-2sz+s^2)} e^{s^2/2} dz \\ &= e^{\frac{s^2}{2}} \underbrace{\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(z-s)^2} dz}_{1}\end{aligned}$$

- The theorem holds because on the right side we have the integral of the Gaussian PDF with mean s and variance 1.

MGF of Gaussian RV (II)

- **Theorem:** If $X \sim N(\mu, \sigma^2)$, then the MGF of X is

$$\phi_X(s) = e^{s\mu + \sigma^2 s^2/2}$$

- **Proof:** $X = \sigma Z + \mu$
  As a property of the MGF, we had seen that the MGF of $Y = aX + b$ satisfied $\phi_Y(s) = e^{sb}\phi_X(as)$. So, in this case, the MGF of would be

$$\phi_X(s) = e^{s\mu}\phi_Z(\sigma s) = e^{s\mu + \sigma^2 s^2/2}$$