Systems and Signals 344

2025

Lecture 23

Moment Generating Functions & Sums of Random Variables

Section 9.1-9.3

Sums of Random Variables

- Random variables of the form $W_n = X_1 + \cdots + X_n$ appear repeatedly in probability theory and applications.
- We could in principle derive the probability model of W_n from the PMF or PDF of $X_1, \ldots, X_n$.
- 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.

Expected Values of Sums

The expected value of a sum of any random variables is the sum of the expected values.

The variance of the sum of any random variable is the sum of all the covariances.

The variance of the sum of independent random variables is the sum of the variances.

REVISION: Expected Value of the sum of two RVs

- The expected value and the variance of the *sum of two random variables* are as follows:

$$\mathrm{E}[X+Y]=\mathrm{E}[X]+\mathrm{E}[Y]$$

- For the expected value, we only need the marginal probability models, and not the complete joint probability models.

$$\mathrm{Var}[X+Y]=\mathrm{Var}[X]+\mathrm{Var}[Y]+2\mathrm{E}[(X-\mu_X)(Y-\mu_Y)]$$

- For the variance we do need the joint probability models.
- The last term is so common it has its own name, the **covariance**.

Expected Values of Sums

For any set of random variables $X_1, \ldots, X_n$, the sum $W_n = X_1 + \cdots + X_n$ has expected value and variance:

$$\mathrm{E}[W_n] = \mathrm{E}[X_1] + \mathrm{E}[X_2] + \cdots + \mathrm{E}[X_n]$$

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

When $X_1, \ldots, X_n$ are uncorrelated

$$\mathrm{Var}[W_n] = \sum_{i=1}^{n} \mathrm{Var}[X_i]$$

Example 9.2

At a party of $n \geq 2$ people, each person throws a hat in a common box. The box is shaken and each person blindly draws a hat from the box without replacement. We say a match occurs if a person draws his own hat. What are the expected value and variance of V_n, the number of matches?

Moment Generating Functions

$\phi_X(s)$, the moment generating function (MGF) of a random variable X, is a probability model of X.

If X is discrete, the MGF is a transform of the PMF.

The MGF of a continuous random variable is a transform of the PDF, similar to a Laplace transform.

The n-th moment of X is the n-th derivative of $\phi_X(s)$ evaluated at $s = 0$.

Moment Generating Functions

- Previously we learned that the PDF of the sum of *independent* random variables ($W_2 = X_1 + X_2$) can be written as the convolution $f_{W_2}(w_2) = f_{X_1}(x_1) * f_{X_2}(x_2)$
- To find the PDF of a sum of three independent random variables, ($W_3 = X_1 + X_2 + X_3$) we first find $f_{W_2}(w_2)$ and then calculate the PDF of $W_3 = W_2 + X_3$ by performing another convolution.
- We could continue this for any n random variables, however, it is generally hard and very time consuming.

Moment Generating Functions

- In linear system theory, *convolution in the time domain* corresponds to *multiplication in the frequency domain* with time functions and frequency functions related by the **Fourier transform**.
- In probability theory, we can, in a similar way, use transform methods to replace the *convolution of PDFs* by *multiplication of transforms*.
- In the language of probability theory, the transform of a PDF or a PMF is a **moment generating function (MGF)**.

We use $\phi_X(s)$ to represent the MGF, in previous years $M_X(t)$ was used.

Moment Generating Functions

For a discrete or continuous random variable X, the **moment generating function (MGF)** of X is

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

Therefore: **Discrete**

$$\phi_X(s) = \sum_{x \in S_X} e^{sx} P_X(x)$$

Continuous

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

Moment Generating Functions

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

- This formulation of the MGF looks very similar to the **Laplace transform** for a time function.
- The primary difference between an MGF and a Laplace transform is that the MGF is defined only for real values of s.
- The set of values of s for which $\phi_X(s)$ exists is called the **region of convergence**.
  - $\phi_X(0) = \mathrm{E}[e^0] = 1$ (always included)
  - X non-negative, then $s \le 0$
  - X bounded so that $P(a < X \le b) = 1$, then $s \in \mathbb{R}$

Moment Generating Functions

A random variable X with MGF $\phi_X(s)$ has n^{th} moment

$$\mathbf{E}[X^n] = m_n = \left.\frac{d^n \phi_X(s)}{ds^n}\right|_{s=0}$$

Note that we only need to determine the MGF once of a distribution and can then simply differentiate, and therefore it is typically easier to calculate the moments of X by finding the MGF and differentiating than by integrating $x^n f_X(x)$.

That is why it is called the moment generating function...

Random Variable	**PMF or PDF**	**MGF $\phi_X(s)$**
Bernoulli (p)	$P_X(x) = \begin{cases} 1-p & x=0 \\ p & x=1 \\ 0 & \text{otherwise} \end{cases}$	$1-p+pe^s$
Binomial (n,p)	$P_X(x) = \binom{n}{x} p^x(1-p)^{n-x}$	$(1-p+pe^s)^n$
Geometric (p)	$P_X(x) = \begin{cases} p(1-p)^{x-1} & x=1,2,\ldots \\ 0 & \text{otherwise} \end{cases}$	$\frac{pe^s}{1-(1-p)e^s}$
Pascal (k,p)	$P_X(x) = \binom{x-1}{k-1} p^k(1-p)^{x-k}$	$\left(\frac{pe^s}{1-(1-p)e^s}\right)^k$
Poisson (α)	$P_X(x) = \begin{cases} \alpha^x e^{-\alpha}/x! & x=0,1,2,\ldots \\ 0 & \text{otherwise} \end{cases}$	$e^{\alpha(e^s-1)}$
Disc. Uniform (k,l)	$P_X(x) = \begin{cases} \frac{1}{l-k+1} & x=k,\ldots,l \\ 0 & \text{otherwise} \end{cases}$	$\frac{e^{sk}-e^{s(l+1)}}{1-e^s}$
Constant (a)	$f_X(x) = \delta(x-a)$	e^{sa}
Uniform (a,b)	$f_X(x) = \begin{cases} \frac{1}{b-a} & a<x<b \\ 0 & \text{otherwise} \end{cases}$	$\frac{e^{bs}-e^{as}}{s(b-a)}$
Exponential (λ)	$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$	$\frac{\lambda}{\lambda-s}$
Erlang (n,λ)	$f_X(x) = \begin{cases} \frac{\lambda^n x^{n-1} e^{-\lambda x}}{(n-1)!} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$	$\left(\frac{\lambda}{\lambda-s}\right)^n$
Gaussian (μ,σ)	$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$	$e^{s\mu+s^2\sigma^2/2}$

Table 9.1 Moment generating function for families of random variables.

Example 9.4

X is an exponential random variable with MGF $\phi_X(s) = \lambda/(\lambda - s)$. What are the first and second moments of X? Write a general expression for the nth moment.

MGF of the Sum of Independent RVs

Moment generating functions are particularly useful for analyzing sums of *independent* random variables, because if X and Y are independent, the MGF of $W = X + Y$ is the product:

$$\phi_W(s) = \mathbf{E}[e^{sX} e^{sY}] = \mathbf{E}[e^{sX}]\mathbf{E}[e^{sY}] = \phi_X(s)\phi_Y(s)$$

Note that this is *only* true if they are independent.

MGF of the Sum of Independent RVs

For a set of independent random variables $X_1, \ldots, X_n$, the moment generating function of $W = X_1 + \cdots + X_n$ is:

$$\phi_W(s) = \phi_{X_1}(s)\phi_{X_2}(s)\ldots\phi_{X_n}(s)$$

When $X_1, \ldots, X_n$ are *iid*, each with MGF $\phi_{X_i}(s) = \phi_X(s)$,

$$\phi_W(s) = [\phi_X(s)]^n$$

Example 9.5

J and K are independent random variables with probability mass functions

$$\begin{array}{c|ccc} j & 1 & 2 & 3 \\ \hline P_J(j) & 0.2 & 0.6 & 0.2 \end{array}, \qquad \begin{array}{c|cc} k & -1 & 1 \\ \hline P_K(k) & 0.5 & 0.5 \end{array}. \tag{9.31}$$

Find the MGF of $M = J + K$. What are $P_M(m)$ and $E[M^3]$?

Joint Moment Generating Function

For a discrete or continuous random variables X and Y, the **moment generating function (MGF)** is

$$\phi_{X,Y}(s_1, s_2) = \mathrm{E}[e^{s_1 X + s_2 Y}]$$

Therefore: **Discrete**

$$\phi_{X,Y}(s_1, s_2) = \sum_{x \in S_X} \sum_{y \in S_Y} e^{s_1 x + s_2 y} P_{X,Y}(x, y)$$

Continuous

$$\phi_{X,Y}(s_1, s_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{s_1 x + s_2 y} f_{X,Y}(x, y)\, dx\, dy$$

Joint Moment Generating Function

A random variables X and Y with MGF $\phi_{X,Y}(s_1, s_2)$ has **joint moment about the origin**

$$\mathrm{E}[X^n Y^k] = m_{n,k} = \frac{\partial^{n+k}}{\partial s_1^n \partial s_2^k} \phi_{X,Y}(s_1, s_2)\Big|_{s_1=s_2=0}$$