Lecture 8

• **Read:** Chapter 6.1-6.6.

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
- Sums of a Random Number of Independent Random Variables

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] and Var[W]
 - 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 cards *n* cards 1, ..., *n*.
- Shuffle and turn over one at a time.
- $X_i = 1$ if the *i*th card is labeled *i*.
- Number of matches is

$$W = X_1 + ... + X_n$$

• Find *E*[*W*].

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

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

• So,
$$E[X_i] = 1/n$$
.

$$E[W] = E[X_1] + ... + E[X_n]$$

= $nE[X_i] = n \cdot 1/n = 1$



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.

$$P[X_2 = 1] = P[X_2 = 1 | 1st \neq label 2] P[1st \neq label 2] + P[X_2 = 1 | 1st = label 2] P[1st = label 2] = \frac{1}{n-1} \cdot \frac{n-1}{n} + 0 \cdot \frac{1}{n} = \frac{1}{n}$$

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 $Cov[X_i, X_i] = 0$ if $i \neq j$.)



Variance of Sums: Example

• X and Y have the joint PDF

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

• 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:

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

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

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

= $2/18 - 2/36 = 1/18$

PDF of the Sum of Two Random Variables

• Theorem: The PDF of W = X + Y is

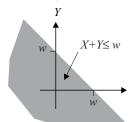
$$f_{W}(w) = \int_{-\infty}^{\infty} f_{X,Y}(x, w - x) dx$$
$$= \int_{-\infty}^{\infty} f_{X,Y}(w - y, y) dy$$

PDF of the Sum of Two Random Variables: Example

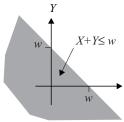
• Find the PDF of W = X + Y when X and Y have the joint PDF

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

- 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 $(0 \le X + Y \le 1)$.



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



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

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

• The complete expression for the PDF of W is

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



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 201E Differential Equations (Laplace transform) (3rd semester)
- TEL 252E Signals and Systems (z-transform, Fourier transform, convolution) (4th semester)

Moment Generating Function (MGF)

• <u>Definition</u>: (Moment Generating Function) For a random variable *X*, the moment generating function (MGF) of *X* is

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

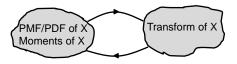
If X is a continuous random variable

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

- 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 $s \le 0$.
- 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

$$\phi_X(s)|_{s=0}=1$$

- 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

$$E[X^n] = \frac{d^n \phi_X(s)}{ds^n} \bigg|_{s=0}$$

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

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

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

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

The MGF of X is

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

MGF of Bernoulli RV

• Let X be a Bernoulli random variable with

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

The MGF of X is

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

MGF of Geometric RV

Let N have a geometric PMF

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

The MGF of N is

$$\phi_{N}(s) = \sum_{n=1}^{\infty} e^{sn} p (1-p)^{n-1}$$

$$= p e^{s} \sum_{n=1}^{\infty} [(1-p)e^{s}]^{n-1}$$

$$= \frac{p e^{s}}{1 - (1-p)e^{s}}$$

MGF of Poisson RV

Let K have the Poisson PMF

$$p_{\mathcal{K}}(k) = egin{cases} rac{lpha^k \mathrm{e}^{-lpha}}{k!} & \text{, } k = 0, 1, ... \\ 0 & \text{, otherwise} \end{cases}$$

The MGF of K is

$$\phi_{\kappa}(s) = \sum_{k=0}^{\infty} e^{sk} \alpha^k e^{-\alpha} / k!$$
$$= e^{-\alpha} \sum_{k=0}^{\infty} (\alpha e^s)^k / k!$$
$$= e^{\alpha (e^s - 1)}$$

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

$$\phi_{Z}(s) = \int_{-\infty}^{\infty} e^{sz} f_{Z}(z) dz$$
$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{sz} e^{-z^{2}/2} dz$$

Completing the square in the exponent:

$$\phi_{Z}(s) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(z^{2} - 2sz + s^{2})} e^{s^{2}/2} dz$$
$$= e^{\frac{s^{2}}{2}} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(z - s)^{2}} dz$$

• 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

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



Finding Moments Using the MGF

 One advantage of the moment generating function is that typically it is easier to find the MGF of X and take derivatives to find the moments of X than it is to find the moments of X directly.

Finding Moments Using the MGF: Example

 Continuing with the example where we had looked at the MGF of the exponential RV, the exponential RV X has first moment

$$E[X] = \frac{d\phi_X(s)}{ds}\bigg|_{s=0} = \frac{\lambda}{(\lambda - s)^2}\bigg|_{s=0} = \frac{1}{\lambda}$$

• The second moment of X is

$$E[X^2] = \frac{d^2 \phi_X(s)}{ds^2} \bigg|_{s=0} = \frac{2\lambda}{(\lambda - s)^3} \bigg|_{s=0} = \frac{2}{\lambda^2}$$

 Proceeding in this way, it should become apparent that the nth moment of X is

$$E[X^n] = \frac{d^n \phi_X(s)}{ds^n} \bigg|_{s=0} = \frac{n! \lambda}{(\lambda - s)^{n+1}} \bigg|_{s=0} = \frac{n!}{\lambda^n}$$

Inversion Property

- $\phi_X(s)$ completely determines the probability law of X.
- Thus, if $\phi_X(s) = \phi_Y(s)$ for all s, then X and Y have the same distribution, i.e., $X \sim Y$.
- Remark: This inversion property is often used in this "pattern matching" mode, i.e., we determine the MGF of an RV and ask "What is the associated distribution?"
 - That is, if you compute and recognize the MGF of an RV, you know its PDF.

Inversion Property: Example 1

• What is the PDF/PMF of X with

$$\phi_X(s) = \frac{1}{4}e^{-s} + \frac{1}{2} + \frac{1}{8}e^{4s} + \frac{1}{8}e^{5s}$$
?

•
$$f_X(x) = \frac{1}{4}\delta(x+1) + \frac{1}{2}\delta(x) + \frac{1}{8}\delta(x-4) + \frac{1}{8}\delta(x-5)$$

Inversion Property: Example 2

• What is the PDF/PMF of X with $\phi_X(s) = e^{10s + 5s^2/2}$?

- If $X \sim N(\mu, \sigma^2)$, $\phi_X(s) = e^{\mu s + s^2 \sigma^2/2}$.
- So, $X \sim N(10, 5)$.

Sum of Two Independent Random Variables

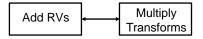
• Assume Z = X + Y and X, Y are independent, i.e.,

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

Calculate transform of the sum

$$\phi_{Z}(s) = E\left[e^{s(X+Y)}\right] = E\left[e^{sX}\right] E\left[e^{sY}\right]$$
$$= \phi_{X}(s)\phi_{Y}(s)$$

So,



convolution of PDFs ⇔ multiplication of MGFs

MGF of the Sum of Independent Random Variables

Theorem: For $X_1, ..., X_n$, a sequence of independent random variables, the moment generating function of $W = X_1 + ... + X_n$ is

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

Proof:

- Convolution of PDFs ⇔ multiplication of MGFs
- When $X_1, ..., X_n$ are <u>iid</u>, $\phi_{X_i}(s) = \phi_X(s)$ for all i and the theorem above has a simple corollary.
- Theorem: For <u>iid</u> $X_1, ..., X_n$, each with MGF $\phi_X(s)$, the moment generating function of $W = X_1 + ... + X_n$ is

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

Sum of Independent Poisson Random Variables

• Theorem: If $K_1, ..., K_n$ are independent Poisson random variables, $W = K_1 + ... + K_n$ is a Poisson random variable.

• Proof:

- $K_1, ..., K_n$ are independent Poisson random variables each with $E[K_i] = \alpha_i$
- What is the MGF of $W = K_1 + ... + K_n$?
- K_i has MGF $\phi_{K_i}(s) = e^{\alpha_i(e^s-1)}$
- Using the theorem on the previous slide,

$$\phi_{W}(s) = e^{\alpha_{1}(e^{s}-1)} e^{\alpha_{2}(e^{s}-1)} ... e^{\alpha_{n}(e^{s}-1)}$$

$$= e^{(\alpha_{1}+\alpha_{2}+...\alpha_{n})(e^{s}-1)}$$

$$= e^{(\alpha_{T})(e^{s}-1)}$$

• W is Poisson with mean $\alpha_T = \alpha_1 + \alpha_2 + ... + \alpha_n$.



Sum of Independent Gaussian Random Variables

• Theorem: The sum of n independent Gaussian random variables $W = X_1 + ... + X_n$ is a Gaussian random variable with mean and variance

$$E[W] = E[X_1] + ... + E[X_n]$$

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

Sum of Independent Random Variables: Special Cases

- Poisson random variables have the special property that the probability model of the sum of n iid random variables has the same form as the probability model of each individual random variable.
- We will see that Gaussian random variables also have this property.
- Except for the Poisson and Gaussian special cases, the probability models of random variables that are sums of iid random variables differ markedly in form from probability models of the random variables in the sum.
 - The sum of iid Bernoulli random variables is a binomial random variable.
 - The sum of iid exponential random variables is an Erlang random variable.

Sum of IID Bernoulli Random Variables

• Theorem: If $X_1, ..., X_n$ are *iid* Bernoulli(p) random variables, then $W = X_1 + ... + X_n$ has the binomial PDF

$$p_W(w) = egin{cases} inom{n}{w} p^w (1-p)^{(n-w)} & \text{, } w=0,1,...,n \\ 0 & \text{, otherwise} \end{cases}$$

Sum of IID Exponential Random Variables

• Theorem: If $X_1, ..., X_n$ are *iid* exponential (λ) random variables, then $W = X_1 + ... + X_n$ has the Erlang PDF

$$f_W(w) = egin{cases} rac{\lambda^n w^{n-1} e^{-\lambda w}}{(n-1)!} & \text{, } w \geq 0 \\ 0 & \text{, otherwise} \end{cases}$$

Sum of a Random Number of Independent Random Variables

• Sum of iid random variables

$$R = X_1 + ... + X_N$$

- The number of terms in the sum (N) is also random variable!
- MGF will be simple when N is independent of $X_1, X_2, ...$

Sum of a Random Number of Independent Random Variables: Examples

• Example 1: K_i = no. of people on bus i, N = no. of buses arriving in 1 hour, R = no. of people arriving in 1 hour

$$R = K_1 + K_2 + ... + K_N$$

Example 2: N= no. of data packets transmitted in one minute. Each packet is successfully received with probability p. Number of successfully received packets in the one-minute span is

$$R = X_1 + X_2 + ... + X_N$$

Note: Since the number N of packets transmitted is random, R is not a binomial random variable!



Sum of a Random Number of Independent Random Variables: Examples (cont.)

- Example 3: The number of children a couple has can be modeled by a discrete random variable X.
 - Suppose in turn all children get married, and have a number of their own children which has the same distribution as X.
 - Find the mean and variance for the number of grandchildren the couple may have.

Sum of a Random Number of Independent Random Variables

• Theorem: Let $\{X_1, X_2, ...\}$ be a collection of <u>iid</u> random variables each with MGF $\phi_X(s)$. Let N be a nonnegative integer-valued random variable that is independent of $\{X_1, X_2, ...\}$. The random sum $R = X_1 + ... + X_N$ has moment generating function

$$\phi_R(s) = \phi_N(\ln \phi_X(s))$$

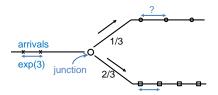
• Theorem: The random sum of <u>iid</u> random variables $R = X_1 + ... + X_N$ has mean and variance

$$E[R] = E[N]E[X]$$

$$Var[R] = E[N]Var[X] + Var[N](E[X])^{2}$$

• Note: Both theorems above require that N be independent of the random sequence $X_1, X_2, ...$. That is, the number of terms in the random sum cannot depend on the actual values of the terms in the sum.

Sum of a Random Number of Independent Random Variables: Example



- Suppose a stream of cars (or data packets), which has independent exponentially distributed (with parameter 3) interarrival times, arrives at an intersection.
- At the intersection, a police officer (or router) flips a coin and with probability 1/3 sends the car (or packet) on the "high road", otherwise sends the car on the "low road" (for load balancing).
- What are the car interarrivals on the high road?

Sum of a Random Number of Independent Random Variables: Example (cont.)

- Let N be the number of arrivals up to and including the first one that gets sent to the high road. ⇒ N ~ geometric(1/3)
- Let X_i , i=1,2,... denote the interarrival times of the cars arriving at the intersection. $\Rightarrow X_i \sim \exp(3)$ and i.i.d.
- When the first car gets sent to the high road, the amount of time equivalent to the sum of the interarrival times up to that point has elapsed. $\Rightarrow X_1 + \cdots + X_N$
- The interarrival times on the high road have a PDF corresponding to $R = X_1 + \cdots + X_N$

Sum of a Random Number of Independent Random Variables: Example (cont.)

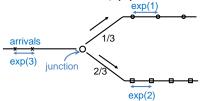
Using the MGFs for geometric and exponential RVs, we have

$$\Phi_N(s) = \frac{e^s/3}{1 - 2e^s/3}$$
 , $\Phi_X(s) = \frac{3}{3 - s}$

and

$$\Phi_R(s) = \Phi_N(ln(\Phi_X(s))) = \Phi_N\left(ln\left(\frac{3}{3-s}\right)\right) = \frac{1}{1-s}$$

 Notice that this transform corresponds to an exponential RV with parameter 1, i.e., R ~ exp(1).



Some Applications for Sums of Random Variables

Study of random walks

```
(A nice applet: http://www.chem.uoa.gr/Applets/AppletSailor/Appl_Sailor2.html)
```

- path of a drunken sailor
- Study of branching processes

```
(A nice applet: http://mcs.open.ac.uk/crj3/Branching/branching.html)
```

- estimating evolutions of multiple generations of families
- probability of extinction?
- Large deviations
 - estimating probabilities of rare events (such as buffer overflow in queueing networks)