SSIE 660: Applied Stochastic Processes
Dr. Sung H. Chung
Note 5
Chapter 2. Random Variables, Chapter 3. Conditional Probability/Expectation

2.7 Moment Generating Functions

Summary

Table 2.1

Discrete probability distribution	Probability mass function, $p(x)$	Moment generating function, $\phi(t)$	Mean	Variance
Binomial with parameters n, p, $0 \le p \le 1$	$\binom{n}{x} p^x (1-p)^{n-x}$, $x = 0, 1, \ldots, n$	$(pe^t + (1-p))^n$	np	$np(1-p)$
Poisson with parameter $\lambda > 0$	$e^{-\lambda} \frac{\lambda^x}{x!}$, $x = 0, 1, 2, \ldots$	$\exp\{\lambda(e^t - 1)\}$	λ	λ
Geometric with parameter $0 \le p \le 1$	$p(1-p)^{x-1}$, $x = 1, 2, \ldots$	$\dfrac{pe^t}{1-(1-p)e^t}$	$\dfrac{1}{p}$	$\dfrac{1-p}{p^2}$

Table 2.2

Continuous probability distribution	Probability density function, $f(x)$	Moment generating function, $\phi(t)$	Mean	Variance
Uniform over (a, b)	$f(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & \text{otherwise} \end{cases}$	$\dfrac{e^{tb} - e^{ta}}{t(b-a)}$	$\dfrac{a+b}{2}$	$\dfrac{(b-a)^2}{12}$
Exponential with parameter $\lambda > 0$	$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & x < 0 \end{cases}$	$\dfrac{\lambda}{\lambda - t}$	$\dfrac{1}{\lambda}$	$\dfrac{1}{\lambda^2}$
Gamma with parameters (n, λ), $\lambda > 0$	$f(x) = \begin{cases} \frac{\lambda e^{-\lambda x}(\lambda x)^{n-1}}{(n-1)!}, & x \ge 0 \\ 0, & x < 0 \end{cases}$	$\left(\dfrac{\lambda}{\lambda - t}\right)^n$	$\dfrac{n}{\lambda}$	$\dfrac{n}{\lambda^2}$
Normal with parameters (μ, σ^2)	$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \times \exp\{-(x-\mu)^2/2\sigma^2\}$, $-\infty < x < \infty$	$\exp\left\{\mu t + \dfrac{\sigma^2 t^2}{2}\right\}$	μ	σ^2

2.8 Limit Theorems

Proposition 1. *Let X be a non-negative random variable, then for any value $a > 0$,*

$$P[X \geq a] \leq \frac{E[X]}{a}$$

It is called a Markov's Inequality

Proof.

$$E[X] = \int_0^\infty x f(x)\,dx = \int_0^a x f(x)\,dx + \int_a^\infty x f(x)\,dx \geq \int_a^\infty x f(x)\,dx$$

$$\geq \int_a^\infty a f(x)\,dx = a \int_a^\infty f(x)\,dx = a P[X \geq a]$$

$\square$

Proposition 2. *If X is a random variable with mean μ and variance σ^2, then, for any value $k > 0$,*

$$P[|x - \mu| \geq k] \leq \frac{\sigma^2}{k^2}$$

It is called Chebyshev's Inequality.

Proof. $(X - \mu)^2$: non negative r.v. so applying the Markov's inequality

$$P[(X-\mu)^2 \geq k^2] \leq \frac{E[(X-\mu)^2]}{k^2}$$

But since $(X-\mu)^2 \geq k^2$ iff $|X-\mu| \geq k$, the preceding is equivalent to

$$P\{|X-\mu| \geq k\} \leq \sigma^2/k^2$$

$\square$

These inequalities help us to obtain bounds for the probabilities, without knowing the actual probability density function - only the mean and variance need to be known.

Example 3. Suppose we know that the number of items produced in a factory during a week is a random variable with a mean of 500.

1. What can be said about the probability that this week's production will be at least 1000?

$P[X \geqslant 1000] \leq \frac{E[X]}{1000} = \frac{500}{1000} = \frac{1}{2}$

2. If the variance of a week's production is known to equal 100, then what can be said about the probability that this week's production will be between 400 and 600?

$P[\,|X - 500| \geqslant 100\,\} \leq \frac{\sigma^2}{(100)^2} = \frac{100}{100^2} = \frac{1}{100}$

$\therefore P[\,|X-500| \leq 100\,\} \geqslant 1 - \frac{1}{100} = \frac{99}{100}$

Theorem 4 (Strong Law of Large Numbers)**.** *Let $X_1, X_2, ...$ be a sequence of independent random variables having a common distribution, and let $E[X_i] = \mu$. Then with probability 1,*

$$\frac{X_1 + X_2 + \cdots + X_n}{n} \to \mu \quad \text{as } n \to \infty$$

Example 5. Let $X_1, X_2, ...$ be a sequence of independent binomial random variables with parameters n and p. Given that $E[X_i] = np$ for all i, we have by the strong law of large numbers that, with probability 1,

$$\frac{X_1 + X_2 + \cdots + X_m}{m} \to np \text{ as } m \to \infty$$

Theorem 6 (Central Limit Theorem: Linderberg)**.** *Let $X_1, X_2, ...$ be a sequence of independent, (not necessarily) identically distributed random variables. Let $E[X_i] = \mu_i$, Var(X_i) = σ_i^2. Let*

$$S_n = \frac{X_1 + X_2 + \cdots + X_n - \sum_i \mu_i}{\sqrt{\sum_i \sigma_i^2}}$$

then the probability function S_n converges to a standard normal distribution N(0,1) as $n \to \infty$.

Theorem 7 (Central Limit Theorem: Special Case)**.** *Let $X_1, X_2, \ldots$ be a sequence of independent, identically distributed (iid) random variables, each with mean μ and variance σ^2. Then the distribution of*

$$\frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}$$

tends to the standard normal as $n \to \infty$. That is,

$$P\left\{\frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \le a\right\} \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-x^2/2} dx \quad \text{as } n \to \infty$$

Example 8. Let X be a binomial random variable with parameters n and p. Then $X = X_1 + \cdots + X_n$ where X_i is a Bernoulli random variable with a parameter p.

1. Find $E[X_i]$

   p

2. Find $\text{Var}(X_i)$.

   $p - p^2 = p(1-p) = pq$

3. As $n \to \infty$, does X follow N(0,1)? Why?

$$\frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} = \boxed{\frac{X - np}{\sqrt{npq}}} \sim \text{Normal}(0,1)$$

as $n \to \infty$

$$_{40}C_{20}\,(.5)^{20}(.5)^{20} = .12537$$

$np = 20$

$npq = 10 \quad \sqrt{npq} = \sqrt{10}$

Example 9. Let X be the number of times that a fair coin, flipped 40 times, lands heads. Find the probability that $X = 20$. Use the normal approximation and then compare it to the exact solution. In addition, find $P[X > 25]$.

$$P\{19.5 \le x \le 20.5\} = P\left(\frac{19.5-20}{\sqrt{10}} < \frac{X-20}{\sqrt{10}} < \frac{20.5-20}{\sqrt{10}}\right)$$

$$= P\{-.16 < Z < .16\} = .1272$$

Continuity correction.

$$P[X > 25] = P[X \ge 26] \doteq P[X > 25.5]$$

$$= P\left[\frac{X-20}{\sqrt{10}} > \frac{25.5-20}{\sqrt{10}}\right) = P(Z \ge 1.74)$$

$$= 1 - P[Z \le 1.74)$$

$$= 1 - .9591 = .0409$$

Example 10. From an urn containing 10 identical balls numbered 0 to 9, n balls are drawn with replacement. Use the Central Limit Theorem to find the probability that among n numbers chosen, the numbers 5 will appear between $\dfrac{n-3\sqrt{n}}{10}$ and $\dfrac{n+3\sqrt{n}}{10}$ times, if (i) $n = 25$, (ii) $n = 100$.

$$X_i = \begin{cases} 1 & \text{if 5 appears on the } i^{th} \text{ draw} \\ 0 & \text{o.w} \end{cases}$$

$$\sum_{i=1}^{n} X_i$$

$$P\left\{\frac{n-3\sqrt{n}}{10} \le \sum X_i \le \frac{n+3\sqrt{n}}{10}\right\}$$

$$= P\left\{\frac{\frac{n-3\sqrt{n}}{10} - \sum\mu_i}{\sqrt{\sum\sigma_i^2}} \le \underbrace{\frac{\sum X_i - \sum\mu_i}{\sqrt{\sum_i \sigma_i^2}}}_{Z \sim N(0,1)} \le \frac{\frac{n-3\sqrt{n}}{10} + \sum\mu_i}{\sqrt{\sum\sigma_i^2}}\right\}$$

$$\mu_i = E[X_i] = 0\cdot P(X_i=0) + 1\cdot P(X_i=1) = \frac{1}{10}.$$

$$\sigma_i^2 = E[X^2] - (E[X])^2 = \frac{1}{10} - \left(\frac{1}{10}\right)^2 = \frac{10}{100} - \frac{1}{100} = \frac{9}{100}$$

2.9 Stochastic Processes

What is a Stochastic Process

- A stochastic process is a collection of random variables (RVs), usually denoted by
$$X = \{X(t) : t \in T\}$$
- t is a call index, most commonly it denotes times.
- T is called the index set.
- $X(t)$ is a random variable (RV) associated with an index t. We say $X(t)$ is the state of the system at time t.
- If $X(t)$ takes values in a set S for every $t \in T$, then S is called the sample space of X.
- For example, $X(t)$ may represent the total number of customers that have entered a supermarket by time t; or the number of the customers in the super market at time t; or the total amount of sales that have been recorded in the market by time t, etc.

Thus, a stochastic process is a family of random variables that describes the evolution through time of some (physical) processes.

Chapter 3. Conditional Probability and Conditional Expectations

3.1 Introduction

i. We are often interested in calculating probabilities and expectations when some partial information is available. ✓

ii. In calculating a desired probability or expectation it is often extremely useful to first condition on some appropriate random variables.

3.2 The Discrete Case

Conditional probability mass function:
$$P[X = x|Y = y] = P_{X|Y}(x|y) = \frac{P(X=x, Y=y)}{P_Y(y)}$$

Conditional expectation:
$$E[X|Y = y] = \sum_{\text{all } x} x\, P(X=x \mid Y=y)$$

Example 11.

x \ y	0	1	2	Marginal $p_X(x)$
1	1/16	3/16	1/16	5/16
2	2/16	1/16	0	3/16
3	5/16	1/16	2/16	8/16
Marginal $p_Y(y)$	8/16	5/16	3/16	1

Find $E[X|Y=0]$

$$\sum x\, P_{X|Y}(x|y=0)$$

$$= 1 \times \frac{1/16}{8/16} + 2 \cdot \frac{2/16}{8/16} + 3 \cdot \frac{5/16}{8/16} = 5/2$$

Example 12. Suppose that $p(x,y)$, the joint probability mass function of X and Y, is given by

$$p(1,1) = 0.5, p(1,2) = 0.1, p(2,1) = 0.1, p(2,2) = 0.3$$

Calculate the conditional probability mass function of X given that $Y = 1$.

X \ Y	1	2	X
1	.5	.1	.6
2	.1	.3	.4
Y	.6	.4	

$$P(X|Y=1) = \frac{P(X=x, Y=y)}{P_Y(y=1)}$$

$x=1$?
$x=2$?

Example 13. If X and Y are independent Poisson random variables with means λ_1, λ_2, respectively, calculate the conditional expected value of X given that $X + Y = n$.

$$P(x) = \frac{e^{-\lambda}\lambda^k}{k!}$$

$$P[X=k \mid X+Y=n] = \frac{P[X=k, X+Y=n]}{P[X+Y=n]} = \frac{P[X=k)P[Y=n-k]}{P[X+Y=n]}$$

$$\frac{e^{-\lambda_1}\lambda_1^k}{k!} \cdot \frac{e^{-\lambda_2}\lambda_2^{n-k}}{(n-k)!}$$

$$= \frac{e^{-(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)^n}{n!}$$

$$\boxed{P(x) = \frac{e^{-\lambda}\lambda^k}{k!}}$$

$$P[X=k \mid X+Y=n] = \frac{P[X=k, X+Y=n]}{P[X+Y=n]} = \frac{P[X=k]P[Y=n-k]}{P[X+Y=n]}$$

$$= \frac{\frac{e^{-\lambda_1}\lambda_1^k}{k!} \cdot \frac{e^{-\lambda_2}\lambda_2^{n-k}}{(n-k)!}}{\frac{e^{-(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)^n}{n!}}$$

$$= \frac{n!}{(n-k)!(k)!} \cdot \frac{\cancel{e^{-\lambda_1-\lambda_2}}\lambda_1^k \cdot \lambda_2^{n-k}}{\cancel{e^{-(\lambda_1+\lambda_2)}}(\lambda_1+\lambda_2)^n}$$

$$= \binom{n}{k}\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^k\left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{n-k}$$

$$P(X+Y=n) = \sum_{k=0}^{n} P(X=k, Y=n-k) = \sum_{k=0}^{n} P(X=k)P(Y=n-k)$$

$$= \sum_{k=0}^{n} \frac{e^{-\lambda_1}\lambda_1^k}{k!} \frac{e^{-\lambda_2}\lambda_2^{n-k}}{(n-k)!} = e^{-(\lambda_1+\lambda_2)} \sum_{k=0}^{n} \frac{\lambda_1^k \lambda_2^{n-k}}{k!(n-k)!}$$

$$= \frac{e^{-(\lambda_1+\lambda_2)}}{n!} \underline{\sum_{k=0}^{n} \frac{n!}{k!(n-k)!} \lambda_1^k \lambda_2^{n-k}}$$

$$= \frac{e^{-(\lambda_1+\lambda_2)}}{n!} (\lambda_1+\lambda_2)^n$$

$X_1 + X_2$ has a Poisson dist w/ mean $\underline{\lambda_1+\lambda_2}$

3.3 The Continuous Case

Conditional probability density function:

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_y(y)}$$

Conditional expectation:

$$E[X|Y=y] = \int x\, f_{X|Y}(x|y)\, dx$$

Example 14. Given

$$f_{X,Y}(x,y) = \begin{cases} x^2 + xy/3, & 0 < x < 1,\ 0 < y < 2 \\ 0, \text{ elsewhere} \end{cases}$$

Find $E[X|Y=y]$.

$$\int_0^1 x\, f_{X|Y}(x|y)\, dx =$$

$$f_Y(y) = \int_0^1 (x^2 + xy/3)\, dx = \left[\frac{1}{3}x^3 + \frac{x^2}{6}y\right]_0^1 = \frac{1}{3} + \frac{y}{6} = \frac{2+y}{6}$$

$0 < y < 2$

$$\therefore f_{X|Y}(x|y) = \frac{\frac{6x^2+2xy}{\cancel{6}}}{\frac{(2+y)}{\cancel{6}}} = \frac{6x^2+2xy}{2+y}$$

$$\therefore \int_0^1 \frac{6x^3+2x^2y}{2+y}\, dx = \left[\frac{6}{2+y}\cdot\frac{1}{4}x^4 + \frac{2}{2+y}\,\frac{1}{3}x^3y\right]_0^1$$

$$= \frac{3}{2(2+y)} + \frac{2y}{(2+y)3} = \frac{9+4y}{6(2+y)}$$

Example 15. Given

$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{2}ye^{-xy}, & 0 < x < \infty,\ 0 < y < 2 \\ 0, \text{ elsewhere} \end{cases}$$

Find $E[e^{X/2}|Y=1]$.

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)} = y e^{-xy}.$$

$$f_Y(y) = \int_0^\infty \frac{1}{2} y e^{-xy} dx = \left[\frac{1}{2}y \cdot \frac{-1}{y} e^{-xy}\right]_0^\infty = 0 - \left(\frac{1}{2}\cancel{y} \cdot \frac{-1}{\cancel{y}}\right)$$

$$= \frac{1}{2}$$

$$\therefore E[e^{X/2}|y=1] = \int e^{x/2} \cdot e^{-x} = \int_0^\infty e^{-\frac{1}{2}x} dx$$

$$= \left[-2e^{-\frac{1}{2}x}\right]_0^\infty = 0 - (-2) = 2$$