MASSACHUSETTS INSTITUTE OF TECHNOLOGY

6.436J/15.085J Fall 2018

Problem Set 4

Readings:

(a) Notes from Lecture 6 and 7.

(b) [Cinlar] Sections I.4, I.5 and II.2

(c) [GS] Chapter 3

Exercise 1. Let N be a random variable that takes nonnegative integer values. Let $X_1, X_2, \ldots$, be a sequence of i.i.d. discrete random variables that have finite expectation and are independent from N. Use iterated expectations to show that the expected value of $\sum_{i=1}^{N} X_i$ is $\mathbb{E}[N]\mathbb{E}[X_1]$.

Solution:

$$\begin{aligned} E[\sum_{i=1}^{N} X_i] &= E[E[\sum_{i=1}^{N} X_i | N]] \\ &= E[NE[X_1]] \\ &= E[N]E[X_1] \end{aligned}$$

Exercise 2. Let X and Y be binomial with parameters (m, p) and (n, q), respectively.

(a) Show that if X is independent from Y, $m = n$, and $p = q$ then $X + Y$ is binomial. *Hint:* Use the interpretation of the binomial, not algebra.

(b) Does the conclusion of part (a) remain valid if $m \neq n$? If X and Y are not independent? If $p \neq q$?

(c) Show that if X and Y are independent, then

$$\mathbb{P}(X + Y = k) = \sum_{i=-\infty}^{\infty} p_X(i) p_Y(k - i).$$

(d) Use the result from part (c) to find the PMF of $X + Y$ where X and Y are independent Poisson random variables with parameters λ and μ, respectively. *Hint:* The "binomial theorem" states that

$$(a + b)^n = \sum_{i=0}^{n} \binom{n}{i} a^i b^{n-i}.$$

Solution:

(a) X and Y can be constructed as sums of i.i.d. Bernoulli random variables. Thus, as long as $p = q$, then $X + Y$ can also be constructed as a sum of i.i.d. Bernoulli random variables, which means that it has a Binomial distribution.

(b) The conclusion of part (a) remains valid if $m \neq n$ using the same argument. If X and Y are not independent, it doesn't hold. For example, for $X = Y$ with $p \in (0, 1)$ and $n > 1$, we have that $X + Y$ only takes values in the even numbers, and not on odds. Thus, it cannot be binomial. If $p \neq q$, the conclusion also does not hold. For example, if $p = 1$, $q = 1/2$, and $n = m = 2$, we have that $\mathbb{P}(X + Y = 1) = \mathbb{P}(X + Y = 2) > \mathbb{P}(X + Y = 3)$, which cannot happen with a Binomial random variable.

(c) If X and Y are independent, then

$$\begin{aligned}
\mathbb{P}(X + Y = k) &= \sum_{i=-\infty}^{\infty} p_{X|Y}(i|k - i) p_Y(k - i) \\
&= \sum_{i=-\infty}^{\infty} p_X(i) p_Y(k - i).
\end{aligned}$$

(d) If X and Y are independent Poisson random variables with parameters λ and μ, respectively, we have

$$\begin{aligned}
\mathbb{P}(X+Y=k) &= \sum_{i=-\infty}^{\infty} p_X(i) p_Y(k-i) \\
&= e^{-\lambda} e^{-\mu} \sum_{i=0}^{k} \frac{\lambda^i \mu^{k-i}}{i!(k-i)!} \\
&= e^{-(\lambda+\mu)} \frac{1}{k!} \sum_{i=0}^{k} \binom{k}{i} \lambda^i \mu^{k-i} \\
&= e^{-(\lambda+\mu)} \frac{(\lambda+\mu)^k}{k!},
\end{aligned}$$

which is a Poisson random variable with parameter $\lambda + \mu$.

Exercise 3. A 4-sided die has its four faces labeled as a, b, c, d. Each time the die is rolled, the result is a, b, c, or d, with probabilities p_a, p_b, p_c, p_d, respectively. Different rolls are statistically independent. The die is rolled n times. Let N_a and N_b be the number of rolls that resulted in a or b, respectively. Find the covariance of N_a and N_b.

Solution: Let $\{X_k \mid k = 1, \ldots, n\}$ be the i.i.d rolls of the dice. Using indicator functions, the number of rolls resulting in a particular outcome is

$$N_e = \sum_{k=1}^{n} \mathbb{1}\{X_n = e\},$$

with mean

$$E[N_e] = \sum_{k=1}^{n} E\left[\mathbb{1}\{X_n = e\}\right] = np_e.$$

The covariance between N_a and N_b is

$$\begin{aligned}\text{cov}(N_a, N_b) = E&\left[\left(\sum_{i=1}^{n} (\mathbb{1}\{X_i = a\} - p_a)\right)\left(\sum_{j=1}^{n} (\mathbb{1}\{X_j = b\} - p_b)\right)\right] \\ &= \sum_{k=1}^{n} E\left[(\mathbb{1}\{X_k = a\} - p_a)(\mathbb{1}\{X_k = b\} - p_b\})\right] \\ &\quad + \sum_{i=1}^{n} \sum_{j \neq i} E\left[(\mathbb{1}\{X_i = a\} - p_a)\left(\mathbb{1}\{X_j = b\} - p_b\right)\right].\end{aligned}$$

This later term is zero since X_i and X_j are independent for $i \neq j$. For this first term

$$\begin{aligned}&E\left[(\mathbb{1}\{X_k = a\} - p_a)(\mathbb{1}\{X_k = b\} - p_b\})\right] \\ &\quad = E\left[\mathbb{1}\{X_k = a\}\mathbb{1}\{X_k = b\}\right] - p_b E\left[\mathbb{1}\{X_k = a\}\right] - p_a E\left[\mathbb{1}\{X_k = b\}\right] + p_a p_b \\ &\quad = 0 - 2p_a p_b + p_a p_b = -p_a p_b.\end{aligned}$$

Hence

$$\text{cov}(N_a, N_b) = -np_a p_b.$$

Exercise 4. Suppose that X and Y are discrete random variables on $(\Omega, \mathcal{F}, \mathbb{P})$. An elegant way of defining the conditional expectation of Y given X is as a random variable of the form $\phi(X)$ (where ϕ is a measurable function), such that

$$\mathbb{E}[\phi(X)g(X)] = \mathbb{E}[Yg(X)],$$

for all measurable functions g. In this problem, we will prove that this condition defines the conditional expectation uniquely; that is, if we also have

$$\mathbb{E}[\psi(X)g(X)] = \mathbb{E}[Yg(X)],$$

for every measurable function g, then $\phi(X)$ and $\psi(X)$ are almost surely equal, i.e., $\mathbb{P}(\phi(X) = \psi(X)) = 1$.

(a) Prove that the following sets are $\mathcal{F}$-measurable: $\{\phi(X) > \psi(X)\}$ and, for any integer n, $A_n := \{\phi(X) > \psi(X) + 1/n\}$.

(b) Assume the contradiction $\mathbb{P}(\phi(X) = \psi(X)) < 1$ and use $g(x) = \mathbf{1}_{A_n}$ for some appropriate n to show that the conditional expectation is unique.

Solution:

(a) First, since X is discrete, $\phi(X)$ and $\psi(X)$ are random variables, without any further assumptions on ϕ, ψ. This implies that $\{\omega|\phi(X(\omega)) > \psi(X(\omega))\}$ and $\{\omega|\phi(X(\omega)) > \psi(X(\omega))\}$ are $\mathcal{F}$-measurable sets; indeed, since one can always fit a rational number between any two distinct real numbers, we have

$$\{\phi(X) > \psi(X)\} = \bigcup_{q \in Q} \{\phi(X) > q\} \bigcap \{q > \psi(X)\}$$

The set $\{\phi(X) > \psi(X) + a\}$ is also $\mathcal{F}$-measurable for any real a (and of course, for $1/n$) since, similarly,

$$\{\phi(X) > \psi(X) + a\} = \bigcup_{q \in Q} \{\phi(X) > q\} \bigcap \{q - a > \psi(X)\}$$

(b) We now proceed to the proof of uniqueness. Assume by contradiction that $\mathbb{P}(\phi(X) = \psi(X)) < 1$. Without loss of generality, we can assume $\mathbb{P}(\phi(X) < \psi(X)) > 0$. The sets

$$A_n = \{\omega|\phi(X)(\omega) + 1/n < \psi(X)(\omega)\}$$

form an increasing sequence of $\mathcal{F}$-measurable sets and

$$\{\omega|\phi(X)(\omega) < \psi(X)(\omega)\} = \bigcup_{n \geq 1} A_n.$$

By continuity of probability, there exists some n such that $\mathbb{P}(A_n) > 0$.

Now, define $g(x) = \mathbf{1}_{A_n}$. Its expectation is well-defined and by construction,

$$\begin{aligned} E[\phi(X)g(X)] &= E[\phi(X)1_{A_n}] \\ &\le E[(\psi(X) - 1/n)1_{A_n}] \\ &= E[\psi(X)g(X)] - (1/n)P(A_n) \\ &< E[\psi(X)g(X)] \end{aligned}$$

which is a contradiction.

Exercise 5. A machine is refilled each morning with n portions of vanilla and chocolate ice creams each (a total of $2n$ portions). Customers arrive sequentially, each getting one of the ice creams independently with probability $1/2$. Consider the first moment when a customer receives an “out of order” message. Let X be the number of portions of the other type left at this moment, $0 \le X \le n$. Find the distribution of X.

Solution: If the unlucky customer ordered vanilla, then there are $X = m$ portions of the chocolate left, if there were $2n - m$ portions were given earlier, out of which n were vanilla. There are $\binom{2n-m}{n}$ sequences of orders that lead to that scenario, all happening with probability $2^{-(2n-m)}$. Thus, since vanilla and chocolate are interchangeable, we have

$$\mathbb{P}[X = m] = \binom{2n-m}{n} 2^{-(2n-m)}.$$

Exercise 6. Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. (So, μ is a measure, but not necessarily a probability measure.) Let $g : \Omega \to \mathbb{R}$ be a nonnegative measurable function. Let $\{B_i\}$ be a sequence of disjoint measurable sets. Prove that

$$\int_{\cup_i B_i} g \, d\mu = \sum_{i=1}^{\infty} \int_{B_i} g \, d\mu.$$

(Be rigorous!)

Note: As an application, this exercise gives another rich source of probability measures. Namely, take f – a nonnegative measurable function on the real line with $\int_{\mathbb{R}} f(x)dx = 1$ (integral w.r.t. Lebesgue measure), and define a set-function $\mathbb{P}(A) = \int_A f dx$. The exercise shows that $\mathbb{P}(\cdot)$ is a probability measure on $(\mathbb{R}, \mathcal{B})$. Function f is called the probability density function (PDF) of $\mathbb{P}$.

Solution: Let $B = \bigcup_{i=1}^{\infty} B_i$ and

$$g_k = g\mathbb{1}_{\cup_{i=1}^k B_i} = \sum_{i=1}^{k} g\mathbb{1}_{B_i},$$

where the last equality holds since the B_i are disjoint. Moreover, for any arbitrary countable collection of disjoint sets $\{U_n\}$

$$\mathbb{1}_{\bigcup_{n\in\mathbb{N}} U_n} = \sum_{n\in\mathbb{N}} \mathbb{1}_{U_n}.$$

Since g is nonnegative g_k is increasing to $g\mathbb{1}_B$, and therefore,

$$\begin{aligned}
\int_B g\,d\mu &= \int g\mathbb{1}_B\,d\mu \\
&= \int \lim_{k\to\infty} g_k\,d\mu \quad (g_k \nearrow g\mathbb{1}_B) \\
&= \lim_{k\to\infty} \int g_k\,d\mu \quad (\text{ Monotone convergence theorem}) \\
&= \lim_{k\to\infty} \int \sum_{i=1}^{k} g\mathbb{1}_{B_i}\,d\mu \\
&= \lim_{k\to\infty} \sum_{i=1}^{k} \int g\mathbb{1}_{B_i}\,d\mu \quad (\text{Linearity of integration}) \\
&= \sum_{i=1}^{\infty} \int_{B_i} g\,d\mu.
\end{aligned}$$

Exercise 7. ***[Optional, not to be graded]*** Let μ and ν be two finite measures on $(\mathbb{R}, \mathcal{B})$. Show that if

$$\int_{\mathbb{R}} f\,d\mu = \int_{\mathbb{R}} f\,d\nu$$

for all bounded continuous functions f then $\mu = \nu$. (*Hint:* write $\mathbb{1}_{(a,b)}(x)$ as an increasing limit of continuous functions.)

Note: This exercise shows that measure on Borel σ-algebra is uniquely characterized by its values on continuous functions. This is true on $\mathbb{R}$, $\mathbb{R}^n$ and any other topological space. Similar to how it is sufficient to know measures only on intervals $(-\infty, a)$ it is sufficient to consider only a handful of functions (such as all sines and cosines, or all exponents). This will be discussed later.

Solution: Let $a < b \in \mathbb{R}$. Consider the sequence of functions, defined for all $n > \frac{2}{b-a}$ and 0 otherwise,

$$f_n(x) = \begin{cases} 0 & x \le a \\ n(x-a) & a < x < a + \frac{1}{n} \\ 1 & a + \frac{1}{n} \le x \le b - \frac{1}{n} \\ -n(x-b) & b - \frac{1}{n} < x < b \\ 0 & x \ge b \end{cases}.$$

By construction the $\{f_n\}$ are increasing, bounded and continuous, piecewise linear. Moreover, $\lim_{n\to\infty} f_n \nearrow \mathbb{1}_{(a,b)}$, this follows since

$$(a, a + \tfrac{1}{n}), (b - \tfrac{1}{n}, b) \to \emptyset \qquad [a + \frac{1}{n}, b - \frac{1}{n}] \to (a, b).$$

Therefore, by the monotone convergence theorem

$$\begin{aligned} \mu(a,b) &= \int \mathbb{1}_{(a,b)}\, d\mu \\ &= \int \lim_{n\to\infty} f_n\, d\mu \quad (f_n \nearrow \mathbb{1}_{(a,b)} \text{ pointwise}) \\ &= \lim_{n\to\infty} \int f_n\, d\mu \quad \text{(Monotone converge theorem)} \\ &= \lim_{n\to\infty} \int f_n\, d\nu \quad (f_n \text{ bounded and continuous}) \\ &= \int \lim_{n\to\infty} f_n\, d\nu \quad \text{(Monotone convergence theorem)} \\ &= \int \mathbb{1}_{(a,b)}\, d\nu = \nu(a,b). \end{aligned}$$

Similar to question 1 of homework 3, if two measures agree on a generating collection for a σ-algebra they agree on the entire σ-algebra. While the open intervals (a, b) have not been explicitly given as a generating collection for the Borel σ-algebra thus far, observe that any arbitrary closed interval can be written as the intersection of open intervals

$$[a, b] = \bigcap_{n=1}^{\infty} (a - \frac{1}{n}, b + \frac{1}{n}).$$