MASSACHUSETTS INSTITUTE OF TECHNOLOGY

6.436J/15.085J Fall 2018

Problem Set 8

Readings:
Notes from Lecture 14 and 15.
[GS]: Section 4.9, 4.10, 5.7-5.9

Exercise 1. Let $\phi_A(t) = \mathbb{E}[e^{itA}]$ be a characteristic function of r.v. A.

(a) Find $\phi_X(t)$ if X is a Bernoulli(p) random variable.

(b) Suppose that $\phi_{X_n} = \cos(t/2^n)$. What is the distribution of X_n?

(c) Let $X_1, X_2, \ldots$ be independent and let $S_n = X_1 + \cdots + X_n$. Suppose that S_n converges almost surely to some random variable S. Show that $\phi_S(t) = \prod_{i=1}^{\infty} \phi_{X_i}(t)$.

(d) Evaluate the infinite product $\prod_{n=1}^{\infty} \cos(t/2^n)$. *Hint:* Think probabilistically; the answer is a very simple expression.

Solution:

(a)
$$\phi_X(t) = (1-p) + pe^{it}.$$

(b) Since
$$\phi_{X_n} = \frac{e^{it/2^n} + e^{it/2^n}}{2},$$
X_n has to be $1/2^n$ with probability $1/2$, and $-1/2^n$ with probability $1/2$.

(c) The interchange of limit and expectation can be justified by appealing to the dominated convergence theorem, since $|e^{itS}| = 1$ and $|e^{itS_n}| = 1$.

(d) The sum of the random variables X_n approaches a uniform random variable in $[-1, 1]$ almost surely. By part (c), the product of $\cos t/2^n$ is the characteristic function of $U[-1, 1]$, which is
$$\int_{-1}^{1} e^{itx}\frac{1}{2}dx = \frac{e^{it} - e^{-it}}{2it} = \frac{\sin t}{t},$$
so
$$\prod_{n=1}^{\infty} \cos(\frac{t}{2^n}) = \frac{\sin t}{t}.$$

Exercise 2. Let X be a random variable with mean, variance, and moment generating function), denoted by $\mathbb{E}[X]$, $\text{var}(X)$, and $M_X(s)$, respectively. Similarly, let Y be a random variable associated with $\mathbb{E}[Y]$, $\text{var}(Y)$, and $M_Y(s)$. Each part of this problem introduces a new random variable Q, H, G, D. Determine the means and variances of the new random variables, in terms of the means, and variances of X and Y.

(a) $M_Q(s) = [M_X(s)]^5$.

(b) $M_H(s) = [M_X(s)]^3[M_Y(s)]^2$.

(c) $M_G(s) = e^{6s}M_X(s)$.

(d) $M_D(s) = M_X(6s)$.

Solution:

(a) The random variable Q is the sum of 5 independent random variables, each distributed as X. Thus, $\mathbb{E}[Q] = 5\mathbb{E}[X]$, $\text{var}(Q) = 5\text{var}(X)$.

(b) The random variable H is the sum of three independent random variables distributed as X, and another two independent random variables distributed as Y. Thus, $\mathbb{E}[H] = 3\mathbb{E}[X] + 2\mathbb{E}[Y]$, $\text{var}(H) = 3\text{var}(X) + 2\text{var}(Y)$.

(c) Multiplying a transform by e^{sa} corresponds to adding a to a random variable. Thus, $\mathbb{E}[G] = \mathbb{E}[X] + 6$, $\text{var}(G) = \text{var}(X)$.

(d) Replacing s by sa corresponds to replacing a random variable X by aX. Thus, $\mathbb{E}[D] = 6\mathbb{E}[X]$, $\text{var}(D) = 36\text{var}(X)$.

Exercise 3. A random (nonnegative integer) number of people K, enter a restaurant with n tables. Each person is equally likely to sit on any one of the tables, independently of where the others are sitting. Give a formula, in terms of the moment generating function $M_K(\cdot)$, for the expected number of occupied tables (i.e., tables with at least one customer).

Solution: Let D be the number of occupied tables. Let $X_1, \ldots, X_n$ be the respective indicator variables of each table, that is, $X_i = 1$ if there is at least one person at at table i, and $X_i = 0$ otherwise. Note that $D = X_1 + \cdots + X_n$. Thus we have:

$$
\begin{aligned}
\mathbb{E}[D] &= \mathbb{E}[\mathbb{E}[D|K]] \\
&= \mathbb{E}[\mathbb{E}[X_1+\cdots+X_n|K]] \\
&= n\cdot\mathbb{E}[\mathbb{E}[X_i|K] \\
&= n\cdot\mathbb{E}\Big[1-\Big(\frac{n-1}{n}\Big)^K\Big] \\
&= n-n\cdot\mathbb{E}\Big[\Big(\frac{n-1}{n}\Big)^K\Big] \\
(\text{letting } s=\log((n-1)/n)) &= n-n\cdot\mathbb{E}[e^{sK}] \\
&= n-n\cdot M_K\big(\log((n-1)/n)\big).
\end{aligned}
$$

Exercise 4. (Problem 7, Section 4.9, [GS]): Let the vector $X_r, 1\le r\le n$ have a multivariate normal distribution with zero means and covariance matrix $V=(v_{ij})$. Show that, conditional on the event $\sum_{i=1}^n X_r=x$, $X_1\stackrel{d}{=}N(a,b)$, where $a=(\rho s/t)x, b=s^2(1-\rho^2)$ and $s^2=v_{11}, t^2=\sum_{ij}v_{ij}, \rho=\sum_i v_{i1}/(st)$.

Solution: Let $S_n=\sum_{k=1}^n x_k$. Since the mapping $(X_1,X_2,\ldots,X_n)\to(X_1,S_n)$ is a linear mapping, and the family of multivariate normal distributions is closed under linear mappings, we find that (X_1,S_n) is a bivariate normal distribution. Furthermore, $E[X_1]=0$, $E[S_n]=0$, var$(X_1)=v_{1,1}=s^2$, var $(S_n)=\sum_{i,j}E[X_iX_j]=\sum_{i,j}v_{i,j}=t^2$, and cov $(X_1,S_n)=\sum_{k=1}^n v_{1,k}$. It follows from the definitions that the correlation of X_1 and S_n is $\left(\sum_{k=1}^n v_{1,k}\right)/(st)$. The desired result follows from the basic properties of bivariate normals proven in the lecture notes.

Exercise 5. Suppose that for every k, the pair (X_k,Y) has a bivariate normal distribution. Furthermore, suppose that the sequence X_k converges to X, almost surely. Show that (X,Y) has a bivariate normal distribution. *Hint:* First show that if X_k is a sequence of normally distributed random variables which converges to X almost surely, then X has to be normally distributed as well. Then use the "right" definition of the bivariate normal.

Solution: For any $a,b\in\Re$, $aX_k+bY\stackrel{as}{\longrightarrow}aX+bY \;\Rightarrow\; aX_k+bY\stackrel{d}{\to}aX+bY$. Let $Z_k=aX_k+bY$ be a sequence of normal random variables with variance σ_k^2. As X_k and Y have zero mean, their corresponding characteristic functions are

$$\phi_k(t)=\exp\left(-\frac{\sigma_k^2}{2}t^2\right).$$

Suppose, for all t, $\phi_k(t) \to \phi(t)$ for some continuous function $\phi(t)$. Consider the sequence $\phi_k(1) = \exp\left(-\frac{\sigma_k^2}{2}\right)$. If this sequence convergences then the composition with any continuous function converges. In particular, $-2\log\phi_k(1) = \sigma_k^2$ converges.

Suppose $\sigma_k^2 \to \infty$, then

$$\phi(t) = \begin{cases} 1 & t = 0 \\ 0 & \text{else} \end{cases}.$$

By the inversion theorem, the PDF of the resulting random variable is one over the entire real line, but this does not integrate to one, a contradiction. Therefore, the $\sigma_k^2 \to \sigma^2$ for some $\sigma^2 \in [0, \infty)$ and $Z = aX + bY$ is normal, i.e. $aX + bY \sim \mathcal{N}(0, \sigma^2)$. Hence, (X, Y) has a bivariate normal distribution.

Exercise 6. Suppose that $X, Z_1, \ldots, Z_n$ have a multivariate normal distribution, and X has zero mean. Furthermore, suppose that $Z_1, \ldots, Z_n$ are independent. Show that $\mathbb{E}[X \mid Z_1, \ldots, Z_n] = \sum_{i=1}^n \mathbb{E}[X \mid Z_i]$. Is this result true without the multivariate normal example? (Prove or give a counterexample.)

Solution: Solution: Let $Z = (Z_1, \ldots, Z_n)$, the multivariate normal. By the conditional expectation formula for multivariate normals

$$\mathbb{E}[X \mid Z_1, \ldots, Z_n] = V_{XZ} V_{ZZ}^{-1} (Z - \mu_Z)$$

Note that V_{XZ} is a $1 \times n$ row vector, V_{ZZ} is an $n \times n$ matrix, and $Z - \mu_Z$ is an $n \times 1$ column vector. Now, by independence of Z_i, V_{ZZ} is the diagonal matrix whose i'th diagonal entry is $\text{var}(Z_i)$. Thus, we can rewrite the right hand side above as,

$$\begin{aligned} \mathbb{E}[X \mid Z_1, \ldots, Z_n] &= \sum_{i=1}^n V_{XZ_i} (Z_i - \mu_{Z_i}) / \text{var}(Z_i) \\ &= \sum_{i=1}^n E[X \mid Z_i]. \end{aligned}$$

Exercise 7. Let $Y_1, \ldots, Y_n$ be independent N(0,1) random variables, and let $X_j = \sum_{r=1}^n c_{jr} Y_r$, for some constants c_{jr}. Show that

$$\mathbb{E}[X_j \mid X_k] = \left(\frac{\sum_r c_{jr} c_{kr}}{{}_r\, c_{kr}^2} \right) X_k.$$

Solution: If (X, Y) are jointly normal with means μ_X, μ_Y respectively, and $\mathrm{cov}(X, Y) = V_{XY}$, and $\mathrm{var}(Y) = V_{YY}$, then $\mathbb{E}[X|Y] = \mu_X + V_{XY}V_{YY}^{-1}(Y - \mu_Y)$. Note that in our case, since the Y_r are independent with zero mean, the X_i are also zero mean. Then, we have:

$$\begin{aligned}
\mathbb{E}[X_j|X_k] &= \mu_{X_j} + V_{jk}V_{kk}^{-1}(X_k - \mu_2) \\
&= \mathbb{E}[X_j X_k]\mathbb{E}[X_k X_k]^{-1} X_k \\
&= \left(\frac{{}^{n}_{r_1,r_2=1}\, c_{jr_1} c_{kr_2} \mathbb{E}[Y_{r_1} Y_{r_2}]}{\sum_{r=1}^{n} c_{kr}^2 \mathbb{E}[Y_r 2] + \quad {}_{r_1 \neq r_2}\, c_{kr_1} c_{kr_2} \mathbb{E}[Y_{r_1} Y_{r_2}]} \right) X_k \\
&= \left(\frac{\sum_r c_{jr} c_{kr}}{{}_{r}\, c_{kr}^2} \right) X_k.
\end{aligned}$$

Exercise 8. [Optional, not for grade] Let X, Y be i.i.d. with finite second moments. Suppose that $X+Y$ and $X-Y$ are independent. Show that they must be Gaussian. (*Hint:* Derive a second order differential equation on $\phi_X(t)$.)

Solution: Let $\phi(t) = \phi_X(t) = \phi_Y(t)$. Using both independence relations and properties of characteristic functions

$$\begin{aligned}
\phi((a+b)t)\phi((a-b)t) &= \phi_{(a+b)X}(t)\phi_{(a-b)Y}(t) \\
&= \phi_{(a+b)X+(a-b)Y}(t) \quad (X \perp\!\!\!\perp Y) \\
&= \phi_{a(X+Y)+b(X-Y)}(t) \\
&= \phi_{a(X+Y)}(t)\phi_{b(X-Y)}(t) \quad (X+Y \perp\!\!\!\perp X-Y) \\
&= \phi_{aX}(t)\phi_{aY}(t)\phi_{bX}(t)\phi_{-bY}(t) \quad (X \perp\!\!\!\perp Y) \\
&= \phi(at)^2\phi(bt)\phi(-bt).
\end{aligned}$$

In other words, as a and b were arbitrary,

$$\phi(t+u)\phi(t-u) = \phi(t)^2\phi(u)\phi(-u).$$

By assumption X and Y have finite second moment and thus ϕ is twice continuously differentiable. Differentiating both sides with respect to u

$$\phi'(t+u)\phi(t-u) - \phi(t+u)\phi'(t-u) = \phi(t)^2\left[\phi'(u)\phi(-u) - \phi(u)\phi'(-u)\right],$$

and

$$\begin{aligned}
&\phi''(t+u)\phi(t-u) - \phi'(t+u)\phi'(t-u) - \phi'(t+u)\phi'(t-u) + \phi(t+u)\phi''(t-u) \\
&= \phi(t)^2\left[\phi''(u)\phi(-u) - \phi'(u)\phi'(-u) - \phi'(u)\phi'(-u) + \phi(u)\phi''(-u)\right].
\end{aligned}$$

Evaluating at $u = 0$

$$\begin{aligned}
2\left(\phi''(t)\phi(t) - \phi'(t)^2\right) &= \phi(t)^2 2\left(\phi''(0)\phi(0) - \phi'(0)^2\right) \\
&= \phi(t)^2 2\left(\phi''(0) - \phi'(0)^2\right).
\end{aligned}$$

The resulting differential equation is

$$\phi''(t)\phi(t) - \phi'(t)^2 - (\phi''(0) - \phi'(0)^2)\phi(t)^2 = 0.$$

Consider the test function

$$f(t) = e^{ic_1 t + c_2 t^2}.$$

The first and second derivate for f are

$$f'(t) = (ic_1 + 2c_2 t)e^{ic_1 t + c_2 t^2}$$
$$f''(t) = 2c_2 e^{ic_1 t + c_2 t^2} + (ic_1 + 2c_2 t)^2 e^{ic_1 t + c_2 t^2}$$

and evaluating at zero

$$f'(0) = ic_1 \quad f''(0) = 2c_2 - c_1^2.$$

Therefore

$$\begin{aligned} &f''(t)f(t) - f'(t)^2 - \left(f''(0) - f'(0)^2\right) f(t)^2 \\ &\quad = \left(2c_2 e^{ic_1 t + c_2 t^2} + (ic_1 + 2c_2 t)^2 e^{ic_1 t + c_2 t^2}\right) e^{ic_1 t + c_2 t^2} \\ &\qquad - (ic_1 + 2c_2 t)^2 e^{2(ic_1 t + c_2 t^2)} - (2c_2 - c_1^2 + c_1^2) e^{2(ic_1 t + c_2 t^2)} \\ &\quad = \left(2c_2 + (ic_1 + 2c_2 t)^2 - (ic_1 + 2c_2 t)^2 - 2c_2\right) e^{2(ic_1 t + c_2 t^2)} \\ &\quad = 0. \end{aligned}$$

Hence $f(t)$ satisfies the differential equation and

$$\phi(t) = e^{ic_1 t + c_2 t^2}.$$

From properties of characteristic functions

$$\phi'(0) = iE[X] \quad \phi''(0) = -E[X^2] = -\text{var}(X) - E[X]^2.$$

Thus

$$c_1 = E[X] \quad c_2 = -\text{var}(X)/2.$$

Therefore,

$$\phi(t) = e^{iE[X]t - \frac{\text{var}(X)}{2} t^2}$$

and by the inversion theorem X and Y are Gaussion.

Exercise 9. [Optional, not for grade] (Problem 20 in p. 142, Section 4.14 of [GS]): Suppose that X and Y are independent and identically distributed, and not necessarily continuous random variables. Show that $X + Y$ cannot be uniformly distributed on $[0, 1]$.

Solution: Suppose X and Y are i.i.d. and $Z = X + Y$ is uniform on $[0, 1]$. Then $X, Y \in [0, 1/2]$ almost everywhere, or else $Z \notin [0, 1]$ with positive probability. Let $z \in [0, 1/4]$. Then, as $X, Y \geq 0$ a.e.,

$$z = \mathbb{P}(Z \leq z) \leq \mathbb{P}(X \leq z, Y \leq z) = \mathbb{P}(X \leq z)\mathbb{P}(Y \leq z) = \mathbb{P}(X \leq z)^2,$$

and similarly, as $X, Y \leq 1/2$ a.e.,

$$z = \mathbb{P}(Z > 1 - z) \leq \mathbb{P}(X > 1/2 - z, Y > 1/2 - z) \leq \mathbb{P}(X > 1/2 - z)^2.$$

Combining provides

$$\mathbb{P}(X \leq z),\, \mathbb{P}\left(X > 1/2 - z\right) \geq \sqrt{z}. \tag{1}$$

Moreover, noting both the upper and lower a.e. bounds of X and Y

$$\begin{aligned} 2z &= \mathbb{P}\left(1/2 - z < Z \leq 1/2 + z\right) \\ &\geq \mathbb{P}\left(X > 1/2 - z, Y \leq z\right) + \mathbb{P}\left(Y > 1/2 - z, X \leq z\right) \\ &= 2\mathbb{P}\left(X > 1/2 - z\right)\mathbb{P}(X \leq z) \\ &\geq 2\sqrt{z}\sqrt{z} \\ &= 2z. \end{aligned}$$

Thus all inequalities are equalities, and in particular

$$\mathbb{P}\left(X > 1/2\right)\mathbb{P}(X \leq z) = z,$$

and, in conjunction with (1), this implies that (1) holds with equality, namely

$$\mathbb{P}(X \leq z) = \mathbb{P}\left(X > \frac{1}{2} - z\right) = \sqrt{z}. \tag{2}$$

Therefore,

$$\begin{aligned} \mathbb{P}\left(X, Y \leq \frac{1}{8}, X + Y > \frac{1}{8}\right) &\geq \mathbb{P}\left(\frac{1}{16} < X \leq \frac{1}{8}, \frac{1}{16} < Y \leq \frac{1}{8}\right) \\ &= \mathbb{P}\left(\frac{1}{16} < X \leq \frac{1}{8}\right)^2 \\ &= \left(\mathbb{P}\quad \frac{1}{8}\right) - \mathbb{P}\left(\frac{1}{16}\right)\right)^2 \\ &= \left(\sqrt{\frac{1}{8}} - \sqrt{\frac{1}{16}}\right)^2 \\ &> 0. \end{aligned}$$

This provides

$$\begin{aligned}
\frac{1}{8} &= \mathbb{P}\left(Z \le \frac{1}{8}\right) \\
&= \mathbb{P}\left(X \le \frac{1}{8}, Y \le \frac{1}{8}\right) - \mathbb{P}\left(X, Y \le \frac{1}{8}, X + Y > \frac{1}{8}\right) \\
&< \mathbb{P}\left(X \le \frac{1}{8}, Y \le \frac{1}{8}\right) = \frac{1}{8},
\end{aligned}$$

a contradiction.

6.436J / 15.085J Fundamentals of Probability
Fall 2018