Homework 4 solutions: probability inequalities, characteristic functions

EE 325 (DD): Probability and Random Processes, Autumn 2016

Instructor: Animesh Kumar, EE, IIT Bombay

Set-A

1. Assume that X is a continuous r.v. with $\phi_X(t)$ as the characteristic function. Are $\mathrm{Re}[\phi_X(t)]$, $\mathrm{Im}[\phi_X(t)]$, and $|\phi_X(t)|^2$ valid characteristic functions? For a complex number z, $\mathrm{Re}[z]$ and $\mathrm{Im}[z]$ represent its real and imaginary parts.

   Solution:

   Characteristic function of a r.v. always exists with two properties: (a) it is non-vanishing in a region around zero and $\Phi(0) = 1$; (b) it is bounded, i.e., $|\Phi(t)| \leq |\Phi(0)| = 1$. These are the necessary condition for the characteristic function to be valid. However, they are not sufficient.

   To show that a function is not a characteristic function, it is sufficient to show that one of the above properties is violated. However, to show that a function is a characteristic function, a corresponding pdf has to be found out.

   If $f_X(x)$ is the pdf of X and $\Phi_X(t)$ is its characteristic function, we can write write the one-one association as $f_X(x) \longleftrightarrow \Phi_X(t)$.

   First note that $f_X(-x) \longleftrightarrow \Phi_X(-t)$. Since $\mathbb{E}()$ is a linear operator, therefore, $[f_X(x) + f_X(-x)]/2 \longleftrightarrow \mathrm{Re}[\Phi_X(t)]$. Next note that $[f_X(x) + f_X(-x)]/2$ is a valid pdf since it is always positive and integrates to one.

   On the other hand, $\mathrm{Im}[\Phi_X(t)]$ is zero at $t = 0$. Thus it cannot be a characteristic function.

   Finally, by the convolution property, $f_X(x) \star f_X(-x) \longleftrightarrow \Phi_X(t)\Phi_X(-t) = |\Phi_X(t)|^2$. This is because of conjugate symmetry of $\Phi_X(t)$, i.e., $\Phi^*_X(t) = \Phi_X(-t)$. Let Y have the pdf $f_X(-x)$ and Y be independent of X. Then, $X + Y$ has the pdf $f_X(x) \star f_X(-x)$. Therefore, $|\Phi_X(t)|^2$ is a valid characteristic function.

2. Let Y be a zero-mean random variable with variance σ^2. Show the one-sided inequality,

   $$\mathbb{P}(Y \geq a) \leq \frac{\sigma^2}{\sigma^2 + a^2}$$

   for $a > 0$. (Hint: use the fact that $Y \geq a \Leftrightarrow Y + c \geq a + c$, for any $c \in \mathbb{R}$.)

   Solution:

   From the hint in the question, we have for any $a > 0, c \in \mathbb{R}$,

   $$\mathbb{P}(Y \geq a) = \mathbb{P}(Y + c \geq a + c) \leq \mathbb{P}\left((Y + c)^2 \geq (a + c)^2\right).$$

   From the standard Markov inequality, we have

   $$\mathbb{P}\left((Y + c)^2 \geq (a + c)^2\right) \quad \leq \quad \frac{\mathbb{E}\left(Y^2 + c^2 + 2cY\right)}{(a + c)^2} = \frac{\sigma^2 + c^2}{(a + c)^2}. \tag{1}$$

   The last step above uses the fact that $\mathbb{E}(Y) = 0$. Differentiate the RHS of the above expression (1), to find the best c suited to the inequality. Equating the first derivative to zero, we have $ac = \sigma^2$, and check for yourself that the second derivative of this function with respect to c, is positive at $c = \sigma^2/a$. Substituting $c = \sigma^2/a$ in (1), we have the necessary result.

   $$\frac{\sigma^2 + c^2}{(a + c)^2} \;=\; \frac{\sigma^2\left(\frac{a^2+\sigma^2}{a^2}\right)}{a^2 + 2\sigma^2 + (\sigma^4/a^2)} \;=\; \frac{\sigma^2}{\sigma^2 + a^2}.$$

Set-B

1. For this problem you may require the Schwarz inequality. Given any two rv X and Y with finite variances, the Schwarz inequality states that

$$[\mathbb{E}(XY)]^2 \leq \left[\mathbb{E}(X^2)\mathbb{E}(Y^2)\right].$$

For a rv Z which is positive, i.e. $Z \geq 0$, show that

$$\mathbb{P}(Z > a) \geq \frac{(\mathbb{E}(Z) - a)^2}{\mathbb{E}(Z^2)},$$

where $a > 0$ is any arbitrary constant. (**Hint:** think of a rv which converts into a probability upon taking expectations.)

Solution:

Apply Schwarz inequality on Z and $Y = \mathbf{1}(Z > a)$ to get

$$[\mathbb{E}(ZY)]^2 \leq \left[\mathbb{E}(Z^2)\mathbb{E}(Y^2)\right] = \mathbb{P}(Z > a)\mathbb{E}(Z^2). \tag{2}$$

since $\mathbb{E}(Y^2) = 0^2 \times \mathbb{P}(Z \leq a) + 1^2 \times \mathbb{P}(Z > a) = \mathbb{P}(Z > a)$. Finally, $\mathbb{E}(ZY) = \mathbb{E}(Z) - \mathbb{E}(Z(1-Y)) = \mathbb{E}(Z) - \mathbb{E}(Z.\mathbf{1}(Z \leq a)) \geq \mathbb{E}(Z) - a$ since $Z \times \mathbf{1}(Z \leq a) \leq a$. Thus, $[\mathbb{E}(ZY)]^2 \geq (\mathbb{E}(Z) - a)^2$. Putting this result in (2), we get the desired result

$$\mathbb{P}(Z > a) \geq \frac{(\mathbb{E}(Z) - a)^2}{\mathbb{E}(Z^2)}.$$

Typically this result is used with $a = 0$.

2. Assume that X is a continuous random variable with,

$$f_X(x) = \frac{c}{1 + |x|^3}, x \in \mathbb{R}.$$

The constant c is selected such that $\int_{\mathbb{R}} f_X(x)\mathrm{d}x = 1$. Find $g_X(r)$. (Hint: think before you act).

Solution: For this random variable, we have $\mathbb{E}\left(X^2\right) = \infty$, since we will have an integral approximately equal to

$$\mathbb{E}\left(X^2\right) \approx \int_0^\infty \frac{dx}{x} = \infty. \tag{3}$$

At $r = 0$, by definition $g_X(0) = 1$. By linearity of expectations and by expanding the integral e^{rX}, see that $g_X(r) = \mathbb{E}\left(e^{rX}\right) \geq \frac{r^2}{2}\,\mathbb{E}\left(X^2\right)$ for any $r \neq 0$. Hence for any $r \neq 0$, $g_X(r) \to \infty$ and does not exists.

3. Construct examples of distributions for X such that,

   (a) The Markov inequality is tight, i.e., there exists a distribution $F_X(x)$ and a point $a \in \mathbb{R}$ such that $\mathbb{P}(X \geq a) = (\mathbb{E}(X)/a)$.

   (b) The Chebyshev inequality is tight, i.e., there exists a distribution $F_X(x)$ and a point $a > 0, a \in \mathbb{R}$ such that $\mathbb{P}(|X - \mathbb{E}(X)| \geq a) = (\sigma_X^2/a^2)$.

Solution:

   (a) See that the Markov inequality $a\mathbb{P}(X \geq a) \leq \mathbb{E}(X)$, cannot be true for any continuous random variable X, since in the continuous case, we have a decreasing $\mathbb{P}(X \geq x)$ function and the area under the curve $(\mathbb{E}(X))$ is definitely larger than the area of a rectangle inscribed under the curve $\mathbb{P}(X \geq x)$, please see Fig 1.7 (reproduced below) in the text for a visualisation. Define the random variable X for any $a > 0, a \in \mathbb{R}$ with pmf

$$X = \begin{cases} 0 & \text{with probability} = p \\ a & \text{with probability} = 1 - p \end{cases}$$

   Note that $\mathbb{E}(X) = a(1-p) = a\mathbb{P}(X \geq a)$.

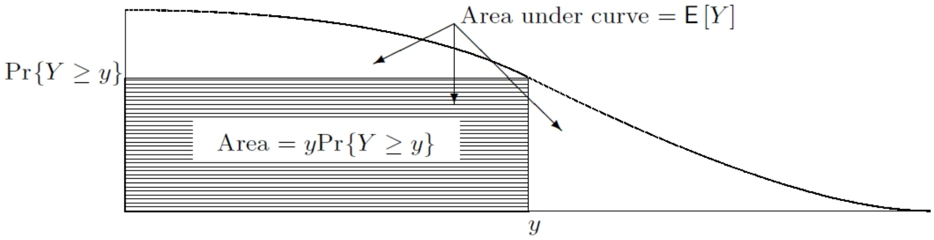

Figure 1: Demonstration that $y\mathbb{P}\left(Y \geq y\right) \leq \mathbb{E}\left(Y\right)$.

(b) Now to define a random variable which is tight for the Chebyshev inequality, modify the above example suitably :). Consider the following discrete random variable X, defined for some $a > 0$.

$$X = \begin{cases} -a & \text{with probability} = p/2 \\ 0 & \text{with probability} = 1-p \\ a & \text{with probability} = p/2 \end{cases}$$

Note that $\mathbb{E}\left(X\right) = 0, \mathbb{E}\left(X^2\right) = a^2 p, \mathrm{Var}(X) = a^2 p$. The Chebyshev inequality is tight for $\mathbb{P}\left(|X| \geq a\right)$.

4. If $Y > 0$ and $\mathbb{E}(Y) < \infty$, then show that $\lim_{y\to\infty} y\mathbb{P}(Y \geq y) = 0$.

Solution:

You are given a random variable which has a finite mean. You are expected to prove that for such a variable $\lim_{y\to\infty} y\mathbb{P}\left(Y \geq y\right) = 0$. Intuitively, consider the question in the following light: $\mathbb{P}\left(Y \geq y\right)$ is a decreasing function of y, and we know that $\mathbb{E}\left(Y\right) = \int_y \mathbb{P}\left(Y \geq y\right) dy$. If $\mathbb{P}\left(Y \geq y\right) = \frac{c}{y}$, for some $c > 0$, then we have

$$\mathbb{E}\left(Y\right) = \int_0^\infty \mathbb{P}\left(Y \geq y\right) dy = c \int_0^\infty \frac{dy}{y} = \infty.$$

$$\mathbb{P}\left(Y \geq y\right) = \frac{1}{yf(y)} \quad \Rightarrow \quad y\mathbb{P}\left(Y \geq y\right) = \frac{1}{f(y)}$$

If this is true, $lim_{y\to\infty}\, y\mathbb{P}\left(Y \geq y\right) = \lim_{y\to\infty} \frac{1}{f(y)} = 0$. Consider a slightly more formal proof as follows, consider the following expression for any $y > 0$:

$$\int_y^\infty u f_Y(u) du \geq y \int_y^\infty f_Y(u) du = y\mathbb{P}\left(Y \geq y\right). \tag{4}$$

You also know the following

$$\begin{aligned} \mathbb{E}\left(Y\right) &= \lim_{z\to\infty} \int_0^z u f_Y(u) du \\ \mathbb{E}\left(Y\right) &= \int_0^y u f_Y(u) du + \int_y^\infty u f_Y(u) du < \infty. \end{aligned} \tag{5}$$

Hence from (4) and (5), we have $\int_y^\infty u f_Y(u) du \to 0$, or $\lim_{y\to\infty} y\mathbb{P}\left(Y \geq y\right) \to 0$.

5. Let $\{X_i\}_{i=1}^\infty$ be an IID sequence of random variables, distributed according to the exponential distribution $\mathrm{Exp}(\lambda)$. Show that,

$$\mathbb{P}\left(\sum_{i=1}^n X_i \geq n\left(\frac{1}{\lambda} + \epsilon\right)\right) \leq \exp\left(n[\ln(1+\lambda\epsilon) - \lambda\epsilon]\right).$$

Show that the bound is non-trivial or the RHS of the inequality is not equal to 1 for $\epsilon > 0$. (Hint: use Chernoff bound formulation.)

Solution:

X, Y are r.v's with non-zero pmf on $\{1, 2, \ldots, m\}$. X_i are i.i.d Exp(λ) r.v's. First calculate the moment generating function $g_{X_1}(r)$.

$$\begin{aligned} g_{X_1}(r) &= \int_0^\infty \lambda e^{-\lambda x} e^{rx} dx \\ &= \frac{\lambda}{\lambda - r} \quad \forall\, r < \lambda. \end{aligned}$$

Since X_i are i.i.d, we have the Chernoff bound as follows

$$\begin{aligned} \mathbb{P}\left(\sum_{i=1}^{n} X_i \geq n\left(\frac{1}{\lambda} + \epsilon\right)\right) &\leq \inf_{r \geq 0} \frac{\mathbb{E}\left(e^{r\sum_{i=1}^n X_i}\right)}{e^{nr\left(\frac{1}{\lambda}+\epsilon\right)}} \\ &= \inf_{r \geq 0} \frac{(g_X(r))^n}{e^{rn\left(\frac{1}{\lambda}+\epsilon\right)}} \\ &= \inf_{r \geq 0}\ e^{-nr\left(\frac{1}{\lambda}+\epsilon\right)} \left(\frac{\lambda}{\lambda - r}\right)^n \\ &= \inf_{r \geq 0}\ e^{-nr\left(\frac{1}{\lambda}+\epsilon\right)\ +\ n\ln\left(\frac{\lambda}{\lambda - r}\right)} \end{aligned} \tag{6}$$

To find the infimum of the above function, first note that the exponent term remains after any number of successive differentiations, and this term is not 0. Hence the differential of the exponent, should be zero and the second differential of this term should be positive to make sure that we have a minima. Equating the differential of the exponent to zero, we have

$$-n\left(\frac{1}{\lambda} + \epsilon\right) + n\ \frac{\lambda - r}{\lambda}\ \frac{\lambda}{(\lambda - t)^2} = 0.$$

Simple algebraic manipulation gives the optimum r, as $r^* = \lambda\left(\frac{\lambda\epsilon}{1+\lambda\epsilon}\right)$. *Check for yourself that the second derivative of the entire RHS in the Chernoff bound expression in* (6) *is indeed positive.* Using this value of r^* in (6), we have the required result.

More justification: It is also easy to see that at $x = 0$, $x = \ln(1 + x)$. For all $x \geq 0$, both the functions $x, \ln(1 + x)$ are increasing. While for any $x > 0$, the derivative of x is 1, while the derivative of $\ln(1 + x)$ is $\frac{1}{1+x} < 1$, which means the function $\ln(1 + x)$ grows slower than linear from the point where it meets x. Hence $\ln(1 + x) \leq x, x \geq 0$ with equality at $x = 0$. Thus see that the exponent is of the form $e^{-n\delta}$, for $\delta = \lambda\epsilon - \ln(1 + \lambda\epsilon) > 0$ which is tight.

6. (Kullback Leibler divergence between discrete random variables:) Let X and Y be two discrete random variables with non-zero pmf defined on the set of integers $\{1, 2, \ldots, m\}$.

   (a) Show that the function $g(x) = \ln x - x + 1 \leq 0$. (Hint: Show that $g''(x) \leq 0$ and find its unique maxima to establish the inequality.)

   (b) Using $\ln x \leq x - 1$ from part (a), show that,

   $$\sum_{i=1}^{m} p_X(i) \ln \frac{p_X(i)}{p_Y(i)} \geq 0, \tag{7}$$

   with equality occuring only if $p_X(i) = p_Y(i)$ for all $i = 1, 2, \ldots, m$. (Hint: Work with the negative of expression on the LHS of the inequality.)

Solution:

(a) If $g(x) = \ln x - x + 1$, see that

$$\begin{aligned} g'(x) &= \frac{1}{x} - 1 \\ g''(x) &= -\frac{1}{x^2}. \end{aligned}$$

Observe that for any $x \in \mathbb{R}$, the second derivative is negative, implying that the function has a maximum at the point of inflection defined at x^*, where $g'(x^*) = 0$. The point of inflection is at $x = 1$, and the maximum value the function takes is $g(1) = 0$. Hence the statement.

(b) From the previous result, we have $\ln x - x + 1 \leq 0$, equivalently, $\ln(1/x) \geq 1 - x$. Now using this result in the first step of the below set of expressions, we have the desired result.

$$\begin{aligned}
\sum_{i=1}^{m} p_X(i)\ \ln \frac{p_X(i)}{p_Y(i)} &= \sum_{i=1}^{m} p_X(i)\ \ln\left(\frac{1}{p_Y(i)/p_X(i)}\right) \\
&\geq \sum_{i=1}^{m} p_X(i)\left(1 - \frac{p_Y(i)}{p_X(i)}\right) \\
&= \sum_{i=1}^{m} p_X(i) - \sum_{i=1}^{m} p_Y(i) = 0.
\end{aligned}$$

You may use Jensen's inequality to prove the same result.