Review Session

Problem 1: Review of Random Variables

Let X and Y be discrete random variables defined on some probability space with a joint pmf $p_{XY}(x, y)$. Let $a, b \in \mathbb{R}$ be fixed.

(a) Prove that $\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$. Do not assume independence.

(b) Prove that if X and Y are independent random variables, then $\mathbb{E}[X \cdot Y] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$.

(c) Assume that X and Y are not independent. Find an example where $\mathbb{E}[X \cdot Y] \neq \mathbb{E}[X] \cdot \mathbb{E}[Y]$, and another example where $\mathbb{E}[X \cdot Y] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$.

(d) Prove that if X and Y are independent, then they are also uncorrelated, i.e.,

$$Cov(X, Y) := \mathbb{E}\left[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])\right] = 0. \tag{1}$$

(e) Find an example where X and Y are uncorrelated but dependent.

(f) Assume that X and Y are uncorrelated and let σ_X^2 and σ_Y^2 be the variances of X and Y, respectively. Find the variance of $aX + bY$ and express it in terms of $\sigma_X^2, \sigma_Y^2, a, b$.
Hint: First show that $Cov(X, Y) = \mathbb{E}[X \cdot Y] - \mathbb{E}[X] \cdot \mathbb{E}[Y]$.

Solution

(a)

$$\begin{aligned}
\mathbb{E}[aX + bY] &= \sum_x \sum_y (ax + by) p_{XY}(x, y) \\
&= \sum_x ax \sum_y p_{XY}(x, y) + \sum_y by \sum_x p_{XY}(x, y) \\
&= a \sum_x x\, p_X(x) + b \sum_y y\, p_Y(y) \\
&= a\mathbb{E}[X] + b\mathbb{E}[Y].
\end{aligned}$$

(b) If X and Y are independent, we have $p_{XY}(x, y) = p_X(x) p_Y(y)$, then

$$\begin{aligned}
\mathbb{E}[X \cdot Y] &= \sum_X \sum_Y xy p_{XY}(x, y) \\
&= \sum_X \sum_Y x p_X(x) y p_Y(y) \\
&= \sum_X x p_X(x) \sum_Y y p_Y(y) \\
&= \mathbb{E}[X] \cdot \mathbb{E}[Y]
\end{aligned}$$

(c) For the first example, suppose $Pr(X = 0, Y = 1) = Pr(X = 1, Y = 0) = \frac{1}{2}$, and $Pr(X = 0, Y = 0) = Pr(X = 1, Y = 1) = 0$. X, Y are dependent, and we have $\mathbb{E}[X \cdot Y] = 0$ while $\mathbb{E}[X]\mathbb{E}[Y] = \frac{1}{4}$

For the second example, suppose $Pr(X=-1,Y=0)=Pr(X=0,Y=1)=Pr(X=1,Y=0)=\frac{1}{3}$. X,Y are dependent. Obviously we have $\mathbb{E}[X\cdot Y]=0$, and furthermore $\mathbb{E}[X]=0$, hence $\mathbb{E}[X]\mathbb{E}[Y]=0$.

(d) If X and Y are independent, we have $p_{XY}(x,y)=p_X(x)p_Y(y)$, then

$$\begin{aligned}
\mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])] &= \sum_x\sum_y (x-\mathbb{E}[X])(y-\mathbb{E}[Y])\,p_{XY}(x,y)\\
&= \sum_x\sum_y (x-\mathbb{E}[X])(y-\mathbb{E}[Y])\,p_X(x)p_Y(y)\\
&= \sum_x (x-\mathbb{E}[X])\,p_X(x)\sum_y (y-\mathbb{E}[Y])\,p_Y(y)\\
&= (\mathbb{E}[X]-\mathbb{E}[X])(\mathbb{E}[Y]-\mathbb{E}[Y])=0.
\end{aligned}$$

Thus, X and Y are uncorrelated.

(e) One example where X and Y are uncorrelated but dependent is

$$\mathbb{P}_{XY}(x,y)=\begin{cases}\frac{1}{3} & \text{if } (x,y)\in\{(-1,0),(1,0),(0,1)\},\\ 0 & \text{otherwise.}\end{cases}$$

First, it can be easily checked that $\mathbb{E}[X\cdot Y]=0=\mathbb{E}[X]\cdot\mathbb{E}[Y]$ (note that $\mathbb{E}[X]=0$). Second, X and Y are dependent since $\mathbb{P}_{XY}(1,0)=\frac{1}{3}$ but $\mathbb{P}_X(1)\mathbb{P}_Y(0)=\frac{1}{3}\times\frac{2}{3}$.

(f) First, we have

$$\begin{aligned}
Cov(X,Y) &= \mathbb{E}\left[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])\right]\\
&= \mathbb{E}\left[XY-X\mathbb{E}[Y]-\mathbb{E}[X]Y+\mathbb{E}[X]\mathbb{E}[Y]\right]\\
&= \mathbb{E}[X\cdot Y]-\mathbb{E}[X]\cdot\mathbb{E}[Y].
\end{aligned}$$

Thus, $Cov(X,Y)=0$ if and only if $\mathbb{E}[X\cdot Y]=\mathbb{E}[X]\cdot\mathbb{E}[Y]$.

Then,

$$\begin{aligned}
\sigma^2_{aX+bY} &= \mathbb{E}[aX+bY-\mathbb{E}[aX+bY]]^2\\
&= \mathbb{E}[(aX+bY)^2]-(\mathbb{E}[aX+bY])^2\\
&= a^2\mathbb{E}[X^2]+2ab\mathbb{E}[X\cdot Y]+b^2\mathbb{E}[Y^2]-a^2\mathbb{E}[X]^2-2ab\mathbb{E}[X]\mathbb{E}[Y]-b^2\mathbb{E}[Y]^2\\
&= a^2(\mathbb{E}[X^2]-\mathbb{E}[X]^2)+b^2(\mathbb{E}[Y^2]-\mathbb{E}[Y]^2)\\
&= a^2\sigma_X^2+b^2\sigma_Y^2.
\end{aligned}$$

We remark that since the independence of X and Y implies $Cov(X,Y)=0$, we also have $\sigma^2_{aX+bY}=a^2\sigma_X^2+b^2\sigma_Y^2$ if X and Y are independent.

Problem 2: Review of Gaussian Random Variables

A random variable X with probability density function

$$p_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-m)^2}{2\sigma^2}} \tag{2}$$

is called a *Gaussian* random variable.

(a) Explicitly calculate the mean $\mathbb{E}[X]$, the second moment $\mathbb{E}[X^2]$, and the variance $Var[X]$ of the random variable X.

(b) Let us now consider events of the following kind:

$$\mathbb{P}(X < \alpha). \tag{3}$$

Unfortunately for Gaussian random variables this cannot be calculated in closed form. Instead, we will rewrite it in terms of the standard Q-function:

$$Q(x) \quad = \quad \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}\, du \tag{4}$$

Express $\mathbb{P}(X < \alpha)$ in terms of the Q-function and the parameters m and σ^2 of the Gaussian pdf.

Like we said, the Q-function cannot be calculated in closed form. Therefore, it is important to have *bounds* on the Q-function. In the next 3 subproblems, you derive the most important of these bounds, learning some very general and powerful tools along the way:

(c) Derive the Markov inequality, which says that for any non-negative random variable X and positive a, we have

$$\mathbb{P}(X \geq a) \quad \leq \quad \frac{\mathbb{E}[X]}{a}. \tag{5}$$

(d) Use the Markov inequality to derive the Chernoff bound: the probability that a real random variable Z exceeds b is given by

$$\mathbb{P}(Z \geq b) \leq \mathbb{E}\left[e^{s(Z-b)}\right], \qquad s \geq 0. \tag{6}$$

(e) Use the Chernoff bound to show that

$$Q(x) \leq e^{-\frac{x^2}{2}} \quad \text{for } x \geq 0. \tag{7}$$

Solution

(a) First,

$$\begin{aligned}
\mathbb{E}[X] &= \int_{-\infty}^{\infty} x p_X(x)\, dx \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} x e^{-\frac{(x-m)^2}{2\sigma^2}}\, dx \\
&\overset{(*)}{=} \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} u e^{-\frac{u^2}{2\sigma^2}}\, du + m \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{u^2}{2\sigma^2}}\, du \\
&\overset{(\dagger)}{=} 0 + m \\
&= m,
\end{aligned} \tag{8}$$

where $(*)$ follows by a change of variable $u = x - m$ and $(\dagger)$ follows since the first integrand in (8) is an odd function and the second integrand in (8) is a probability density function. We remark that the integral

$$\int_{-\infty}^{\infty} e^{-x^2}\, dx$$

known as *Gaussian integral*, can be evaluated explicitly to be $\sqrt{\pi}$. Second,

$$\begin{aligned}
\mathbb{E}[X^2] &= \int_{-\infty}^{\infty} x^2 p_X(x)\,dx \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} x^2 e^{-\frac{(x-m)^2}{2\sigma^2}}\,dx \\
&\overset{(*)}{=} \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} u^2 e^{-\frac{u^2}{2\sigma^2}}\,du + \frac{2m}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} u e^{-\frac{u^2}{2\sigma^2}}\,du + m^2 \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{u^2}{2\sigma^2}}\,du \qquad (9) \\
&\overset{(\dagger)}{=} \sigma^2 + 0 + m^2 \\
&= \sigma^2 + m^2,
\end{aligned}$$

where $(*)$ follows by a change of variable $u = x - m$ and $(\dagger)$ follows from the same arguments in the evaluation of $\mathbb{E}[X]$ and an integration by parts to the first integral in (9):

$$\begin{aligned}
\frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} u^2 e^{-\frac{u^2}{2\sigma^2}}\,du &= -\frac{\sigma^2}{\sqrt{2\pi\sigma^2}} \left(u e^{-\frac{u^2}{2\sigma^2}} \Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} e^{-\frac{u^2}{2\sigma^2}}\,du \right) \\
&= 0 + \sigma^2.
\end{aligned}$$

Therefore,

$$\begin{aligned}
Var[X] &= \mathbb{E}[X - \mathbb{E}[X]]^2 \\
&= \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\
&= \sigma^2 + m^2 - m^2 \\
&= \sigma^2.
\end{aligned}$$

(b)

$$\begin{aligned}
\mathbb{P}(X < \alpha) &= \int_{-\infty}^{\alpha} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-m)^2}{2\sigma^2}}\,dx \\
&\overset{(*)}{=} \int_{-\infty}^{\frac{\alpha-m}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}\,du \\
&= 1 - Q\left(\frac{\alpha - m}{\sigma}\right),
\end{aligned}$$

where $(*)$ follows by a change of variable $u = \frac{x-m}{\sigma}$.

(c)

$$\begin{aligned}
\mathbb{E}[X] &= \int_0^a x p_X(x)\,dx + \int_a^{\infty} x p_X(x)\,dx \\
&\geq 0 + a \int_a^{\infty} p_X(x)\,dx \\
&= a\mathbb{P}(X \geq a).
\end{aligned}$$

(d) Fix $s \geq 0$, then we have

$$\begin{aligned}
\mathbb{P}(Z \geq b) &\leq \mathbb{P}(s(Z-b) \geq 0) \\
&= \mathbb{P}(e^{s(Z-b)} \geq e^0) \\
&\overset{(*)}{\leq} \mathbb{E}\big[e^{s(Z-b)}\big],
\end{aligned}$$

where $(*)$ follows from the Markov inequality.

(e) Let X be a Gaussian random variable with mean zero and unit variance, then we have

$$
\begin{aligned}
Q(x) &= \mathbb{P}(X \geq x) \\
&\overset{(*)}{\leq} \mathbb{E}\left[e^{s(X-x)}\right] \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{s(u-x)} e^{-\frac{u^2}{2}} \, du \\
&= e^{-sx+\frac{s^2}{2}} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(u-s)^2}{2}} \, du \\
&= e^{-sx+\frac{s^2}{2}},
\end{aligned}
$$

where $(*)$ follows from the Chernoff bound. In order to get the tightest bound, we need to minimize $-sx + s^2/2$ which gives $s = x$ and then the desired bound is established.

Problem 3: Moment Generating Function

Let X be a real-valued random variable taking values on a finite set. The logarithmic moment generating function is defined as follows.

$$\phi(s) := \ln E[\exp(sX)] = \ln \sum_x p(x) \exp(sx)$$

(a) Show that $p_s(x) := p(x)\exp(sx)\exp(-\phi(s))$ is a probability mass function.

(b) Let X_s be a random variable taking the same value as X but with probabilities $p_s(x)$, show that $\phi'(s) = E[X_s]$.

(c) Show that
$$\phi''(s) = \mathrm{Var}(X_s) := E[X_s^2] - E[X_s]^2$$
and conclude that $\phi''(s) \geq 0$ and the inequality is strict except when X is deterministic.

(d) Let $x_{\min} := \min\{x : p(x) > 0\}$ and $x_{\max} := \max\{x : p(x) > 0\}$ be the smallest and largest values X takes. Show that
$$\lim_{s \to -\infty} \phi'(s) = x_{\min}, \quad \text{and} \quad \lim_{s \to \infty} \phi'(s) = x_{\max}.$$

Solution

(a) To show that $p_s(x)$ is a probability mass function, we need to show (i) $p_s(x)$ is non-negative for all x and (ii) $\sum_x p_s(x) = 1$. For (i), since $p(x) \geq 0$, $\exp(sx) \geq 0$ and $\exp(-\phi(s)) \geq 0$, we have $p_s(x) \geq 0$. For (ii), the sum of all $p_s(x)$ can be computed as follows.

$$
\begin{aligned}
\sum_x p_s(x) &= \sum_x p(x)\exp(sx)\exp(-\phi(s)) && (10) \\
&= \exp(-\phi(s)) \sum_x p(x)\exp(sx) && (11) \\
&= \exp(-\ln E[\exp(sX)])E[\exp(sX)] && (12) \\
&= \frac{E[\exp(sX)]}{E[\exp(sX)]} = 1 && (13)
\end{aligned}
$$

Both conditions are satisfied. Hence, $p_s(x)$ is a probability mass function.

(b,c) As $\phi(s) := \ln E[\exp(sX)]$, we have

$$\phi'(s) = \frac{E[X\exp(sX)]}{E[\exp(sX)]} = E[X\exp(sX)\exp(-\phi(s))] = E[X_s] \tag{14}$$

$$\phi''(s) = \frac{E[X^2\exp(sX)]}{E[\exp(sX)]} - \frac{E[X\exp(sX)]E[X\exp(sX)]}{E[\exp(sX)]^2} \tag{15}$$

The second term is $E[X_s]^2$ and the first term equals $\sum_x x^2\exp(sx)/\exp(\phi(s)) = E[X_s^2]$. So $\phi''(s) = \mathrm{Var}(X_s)$. Moreover, $\mathrm{Var}(X_s) \geq 0$ with equality only when X_s is deterministic. But X_s is deterministic only when X is.

(d) Observe that

$$\begin{aligned}\phi'(s) &= \frac{E[X\exp(sX)]}{E[\exp(sX)]} = \frac{E[X\exp(sX)]\exp(-sx_{max})}{E[\exp(sX)]\exp(-sx_{max})} && (16)\\ &= \frac{\sum_x p(x)x\exp(-s(x_{max}-x))}{\sum_x p(x)\exp(-s(x_{max}-x))} && (17)\end{aligned}$$

In the sums above, as $s \to \infty$, all terms vanish except the ones for $x = x_{max}$. Hence we have

$$\lim_{s\to\infty}\phi'(s) = \frac{p(x_{max})x_{max}}{p(x_{max})} = x_{max} \tag{18}$$

Similarly, we can show that $\lim_{s\to-\infty}\phi'(s) = x_{min}$.