UC Berkeley
Department of Electrical Engineering and Computer Sciences

EE126: Probability and Random Processes

Problem Set 8 Solutions
Spring 2011

Solution to Problem 1. (a) Let K be the random variable for the number of guards that you bumped into on your way. K has a binomial distribution with parameters $n = 78$ and $p = 1/2$. Hence,

$$p_K(k) = \begin{cases} \binom{78}{k}\left(\frac{1}{2}\right)^k\left(\frac{1}{2}\right)^{78-k} & k = 0, 1, \ldots, 78 \\ 0 & \text{otherwise} \end{cases}$$

Given that you bumped into k guards, random variable X is the sum of k independent normal random variables, each with mean 1 and standard deviation of 1/2. Therefore, conditioned on $K = k$, X is is a normal random variable with mean k and standard deviation $(1/2)\sqrt{k}$:

$$f_{X|K}(x|k) = \frac{1}{\sqrt{2\pi}\frac{1}{2}\sqrt{k}} e^{-\frac{(x-k)^2}{\frac{1}{2}k}} \tag{1}$$

$$= \sqrt{\frac{2}{\pi k}} e^{-\frac{2(x-k)^2}{k}} \tag{2}$$

The transform of X conditioned on $K = k$ is therefore

$$E[e^{sX}|K = k] = e^{(ks^2/8)+ks} \tag{3}$$

Using the total probability theorem, we get:

$$f_X(x) = \sum_{k=0}^{78} p_K(k) f_{X|K}(x|k) \tag{4}$$

$$= \sum_{k=0}^{78} \binom{78}{k}\left(\frac{1}{2}\right)^k\left(\frac{1}{2}\right)^{78-k}\sqrt{\frac{2}{\pi k}} e^{-\frac{2(x-k)^2}{k}} \tag{5}$$

Similarly, since we know the conditional transform of X given $K = k$, we can use the total expectation theorem to get

$$M_X(s) = \sum_{k=0}^{78} p_K(k) E[e^{sX}|K = k] \tag{6}$$

$$= \sum_{k=0}^{78} \binom{78}{k}\left(\frac{1}{2}\right)^k\left(\frac{1}{2}\right)^{78-k} e^{(ks^2/8)+ks} \tag{7}$$

Thus X is a mixture of normal random variables, and its transform is a mixture of the corresponding normal transforms. Note, however, that X itself is not normal!

(b) Let K be the number of gaurds that you bump into. We can view X as the sum of K independent normal random variables, each with mean 1 and standard deviation of 1/2. Thus the transform associated with X can be found by replacing in the binomial transform $M_K(s) = (\frac{1}{2} + \frac{1}{2}e^s)^{78}$ the occurrences of e^s by the normal transform corresponding to $\mu = 1$ and $\sigma = \frac{1}{2}$. Thus,

$$M_X(s) = \left(\frac{1}{2} + \frac{1}{2}\left(e^{\frac{s^2}{8}+s}\right)\right)^{78}.$$

Solution to Problem 2. (a) First, by the tower property $\mathbf{E}[X|Y]$ has mean zero. We write $X = (X - \mathbf{E}[X|Y]) + \mathbf{E}[X|Y]$ to get

$$\mathbf{E}[X\mathbf{E}[X|Y]] = \mathbf{E}[(X - \mathbf{E}[X|Y])\mathbf{E}[X|Y]] + \mathbf{E}[\mathbf{E}[X|Y]^2] \geq 0\,.$$

Here we used the fact that $(X - \mathbf{E}[X|Y])$ and X are uncorrelated.

(b) We first consider the correlation of Y and $\mathbf{E}[X|Y]$: by the tower property

$$\mathbf{E}[Y\mathbf{E}[X|Y]] = \mathbf{E}[\mathbf{E}[YX|Y]] = \mathbf{E}[XY]\,.$$

So we see that in fact the correlation of Y and $\mathbf{E}[X|Y]$ is equal to the correlation of X and Y and so the correlation coefficient must have the same sign.

Solution to Problem 3. ...
In this problem, we want to find the linear estimator $g(Y_1, Y_2)$ that minimizes $\mathbf{E}[(X - g(Y_1, Y_2))^2]$. This is an extension of the case of single measurement. Our estimator is of the form

$$g(Y_1, Y_2) = a_1Y_1 + a_2Y_2 + b$$

and therefore our goal is to find a_1, a_2 and b that solve

$$\text{minimize } (\mathbf{E}[(X - a_1Y_1 - a_2Y_2 - b)^2]) \qquad (*)$$

To achieve $(*)$, we must satisfy

$$\begin{aligned}
\frac{\partial \mathbf{E}[(X-a_1Y_1-a_2Y_2-b)^2]}{\partial b} &= \mathbf{E}[2(X - a_1Y_1 - a_2Y_2 - b)(-1)] &= 0\\
\frac{\partial \mathbf{E}[(X-a_1Y_1-a_2Y_2-b)^2]}{\partial a_1} &= \mathbf{E}[2(X - a_1Y_1 - a_2Y_2 - b)(-Y_1)] &= 0\\
\frac{\partial \mathbf{E}[(X-a_1Y_1-a_2Y_2-b)^2]}{\partial a_2} &= \mathbf{E}[2(X - a_1Y_1 - a_2Y_2 - b)(-Y_2)] &= 0
\end{aligned}$$

and $\partial^2\mathbf{E}[\cdot]/\partial b^2$, $\partial^2\mathbf{E}[\cdot]/\partial {a_1}^2$, and $\partial^2\mathbf{E}[\cdot]/\partial {a_2}^2$ must be > 0.
By the linearity of expectation,

$$\begin{aligned}
b &= \mathbf{E}[X] - a_1\mathbf{E}[Y_1] - a_2\mathbf{E}[Y_2] && (1)\\
a_1\mathbf{E}[Y_1^2] &= \mathbf{E}[XY_1] - a_2\mathbf{E}[Y_1Y_2] - b\mathbf{E}[Y_1] && (2)\\
a_2\mathbf{E}[Y_2^2] &= \mathbf{E}[XY_2] - a_1\mathbf{E}[Y_1Y_2] - b\mathbf{E}[Y_2]. && (3)
\end{aligned}$$

We now have 3 equations to solve for 3 unknowns.
Consider (2) $-$ $\mathbf{E}[Y_1]\cdot(1)$, we obtain

$$a_1\mathbf{E}[Y_1^2]-b\mathbf{E}[Y_1]=\mathbf{E}[XY_1]-a_2\mathbf{E}[Y_1Y_2]-b\mathbf{E}[Y_1]-\mathbf{E}[X]\mathbf{E}[Y_1]+a_1(\mathbf{E}[Y_1])^2+a_2\mathbf{E}[Y_1]\mathbf{E}[Y_2].$$

Arranging algebra and use the fact that $\mathbf{E}[Y_1Y_2]=\mathbf{E}[Y_1]\mathbf{E}[Y_2]$

$$a_1(\mathbf{E}[Y_1^2]-(\mathbf{E}[Y_1])^2)=\mathbf{E}[XY_1]-\mathbf{E}[X]\mathbf{E}[Y_1].$$

Similarly,

$$a_2(\mathbf{E}[Y_2^2]-(\mathbf{E}[Y_2])^2)=\mathbf{E}[XY_2]-\mathbf{E}[X]\mathbf{E}[Y_2].$$

Therefore,

$$\begin{aligned}a_1 &= (\mathbf{E}[XY_1]-\mathbf{E}[X]\mathbf{E}[Y_1])/\sigma_{Y_1}^2\\ a_2 &= (\mathbf{E}[XY_2]-\mathbf{E}[X]\mathbf{E}[Y_2])/\sigma_{Y_2}^2\\ b &= \mathbf{E}[X]-a_1\mathbf{E}[Y_1]-a_2\mathbf{E}[Y_2].\end{aligned}$$

Writing this expression in the similar term as the case of single measurement:

$$g(Y_1,Y_2)=\mathbf{E}[X]+\frac{\mathrm{Cov}(X,Y_1)}{\sigma_{Y_1}^2}(Y_1-\mathbf{E}[Y_1])+\frac{\mathrm{Cov}(X,Y_2)}{\sigma_{Y_2}^2}(Y_2-\mathbf{E}[Y_2])$$

where we use the fact that $\mathrm{Cov}(X,Y)=\mathbf{E}[(X-\mathbf{E}[X])(Y-\mathbf{E}[Y])]=\mathbf{E}[XY]-\mathbf{E}[X]\mathbf{E}[Y]$.
Note: Convince yourself that the second order condition is satisfied.

Solution to Problem 4. To begin, note that $\mathbf{E}S_n=n/2$ and $\mathrm{Var}(S_n)=n/4$. The Central Limit Theorem tells us that $\frac{S_n-n/2}{\sqrt{n}/2}\to X$ in distribution, where $X\sim\mathcal{N}(0,1)$ denotes a standard normal r.v.

(a) We have

$$\begin{aligned}\mathbf{P}(\frac{n}{2}-10\le S_n\le\frac{n}{2}+10) &= \mathbf{P}(\frac{-10}{\sqrt{n}/2}\le\frac{S_n-n/2}{\sqrt{n}/2}\le\frac{10}{\sqrt{n}/2})\\ &\approx \mathbf{P}(X\in[-20/\sqrt{n},20/\sqrt{n}])\to 0\,.\end{aligned}$$

Of course, heuristically it becomes very unlikely that the number of successes will be within 10 of the mean as $n\to\infty$.

(b)

$$\begin{aligned}\mathbf{P}(\frac{n}{2}-\frac{n}{10}\le S_n\le\frac{n}{2}+\frac{n}{10}) &= \mathbf{P}(\frac{-n}{10\sqrt{n}/2}\le\frac{S_n-n/2}{\sqrt{n}/2}\le\frac{n}{10\sqrt{n}/2})\\ &\approx \mathbf{P}(X\in[-\frac{1}{5}\sqrt{n},\frac{1}{5}\sqrt{n}])\to 1\,.\end{aligned}$$

(c)

$$\begin{aligned}\mathbf{P}(\frac{n}{2}-\frac{\sqrt{n}}{2}\le S_n\le\frac{n}{2}+\frac{\sqrt{n}}{2}) &= \mathbf{P}(\frac{-\sqrt{n}}{10\sqrt{n}/2}\le\frac{S_n-n/2}{\sqrt{n}/2}\le\frac{\sqrt{n}}{10\sqrt{n}/2})\\ &\to \mathbf{P}(X\in[-1/5,1/5])=\Phi(.2)-\Phi(-.2)\approx 0.159\,.\end{aligned}$$

Solution to Problem 5. In this problem we will see that Chebyshev's inequality is indeed tight for some random variables. To begin, we might as well assume that $\mathbf{E}X = \mu = 0$ since shifting by μ won't change the probabilities involved here. Now, given σ and c we want to make the probability that $|X| \geq c$ as big as possible while keeping the variance equal to σ. To do this we simply let $\mathbf{P}(X = c) = \mathbf{P}(X = -c) = a/2$ and $\mathbf{P}(X = 0) = 1 - a$ for some value a to be determined shortly. Then we get $\sigma^2 = \text{Var}X = \mathbf{E}[X^2] = ac^2$, so we must put $a = \sigma^2/c^2$. With these definitions we see that $\mathbf{P}(|X - \mu| \geq c) = a = \sigma^2/c^2$.

Solution to Problem 6. (a) When using just one mold, the length of the path is $25X$ and the desired probability is

$$\mathbf{P}(|25X - 1000| < 7.5) \;=\; \mathbf{P}(|X - 40| < 0.3) \;=\; \frac{\sqrt{3}}{10} \;\approx\; 0.1732.$$

(b) When using separate molds with lengths $X_1, X_2, \ldots, X_{25}$ the desired probability is

$$\begin{aligned}\mathbf{P}\left(\left|\left(\sum_{i=1}^{25} X_i\right) - 1000\right| < 7.5\right) \;&=\; \mathbf{P}\left(\frac{|(\sum_{i=1}^{25} X_i) - 1000|}{\sqrt{25}} < \frac{7.5}{\sqrt{25}}\right)\\ &\approx\; \mathbf{P}(|Z| < 1.5) \qquad \text{where } Z \text{ is a standard normal r.v. (CLT)}\\ &=\; \Phi(1.5) - \Phi(-1.5) \;=\; 2\Phi(1.5) - 1 \approx 0.8664.\end{aligned}$$

Intuitively, adding independent instances of mold lengths "averages out" the variations and gives higher probability of a total path length close to the mean.

Solution to Problem 7. We need to show that Y_n converges to $1/2$ in probability, i.e. for each $\epsilon > 0$ we have

$$\mathbf{P}(|Y_n - 1/2| > \epsilon) \to 0\,.$$

Note that

$$\mathbf{P}(|Y_n - 1/2| > \epsilon) = \mathbf{P}(Y_n < 1/2 - \epsilon) + \mathbf{P}(Y_n > 1/2 + \epsilon) = 2\mathbf{P}(Y_n < 1/2 - \epsilon)$$

by symmetry. Let Z_i denote the indicator function for $X_i < 1/2 - \epsilon$, and $S_n = \sum_{i=1}^{2n+1} Z_i$. Note that $\mathbf{E}S_n = (2n+1)(1/2 - \epsilon)$. Now

$$\{Y_n < 1/2 - \epsilon\} \subseteq \{S_n \geq n\}\,.$$

(Think about this carefully! This means there must have been at least n values less than $1/2 - \epsilon$.) But

$$\mathbf{P}(S_n \geq n) \leq \mathbf{P}\left(\frac{S_n - (2n+1)(1/2 - \epsilon)}{2n+1} \geq \epsilon/2\right) \to 0$$

by the weak law of large numbers. This implies that $\mathbf{P}(Y_n < 1/2 - \epsilon) \to 0$, which gives the desired result.

Solution to Problem 8. (a) Let X be a random variable with $\mathbf{P}(X = i) = \frac{1}{i^3}$ for each $i \geq 1$. Then $\mathbf{E}X = \sum \frac{1}{i^2} < \infty$ but $\mathbf{E}X^2 = \sum \frac{1}{i} = \infty$ and hence $\mathrm{Var}X = \infty$. This implies also that $\mathrm{Var}\bar{X}_n = \infty$. This means that $\bar{X}_n$ is very spread out – but still the weak law says $\bar{X}_n$ is approximately equal to the value $\mathbf{E}X_1$ in a probabilistic sense.

(b) We have $\mathrm{Var}Y_i^M \leq \mathbf{E}(Y_i^M)^2 \leq M^2$, which implies that $\mathrm{Var}\bar{Y}_n^M \leq \frac{1}{n}M^2$.