Homework 2

1. ***Predicting outcomes***

Consider the pmf

$$p = \frac{1}{1000}\begin{bmatrix}7 & 9 & 10 & 36 & 41 & 43 & 55 & 73 & 72 & 71 & 72 & 73 & 72 & 72 & 72 & 72 & 73 & 77\end{bmatrix}$$

on $\Omega = \{1, \ldots, 18\}$. Let x be the identity random variable, i.e., $x : \Omega \to \Omega$ defined by $x(\omega) = \omega$.

(a) Compute the mean and covariance of x.

(b) We would like to pick an *estimate* of x. Find the estimate $\hat{x}_{\text{abs}} \in \Omega$ that minimizes the cost function

$$\mathrm{E}|x - \hat{x}|$$

Also find the minimum possible cost.

(c) Find the estimate $\hat{x}_{\text{min-error}} \in \Omega$ that minimizes the probability of error

$$\mathrm{Prob}(x \neq \hat{x})$$

Also find the minimum probability of error.

(d) Find the estimate $\hat{x}_{\text{squared}} \in \Omega$ that minimizes the cost function

$$\mathrm{E}\big((x - \hat{x})^2\big)$$

What is the minimum cost? Show that, if we relax the constraint that $\hat{x}$ lie in Ω, the $\hat{x}$ that minimizes this cost function is

$$\hat{x} = \mathrm{E}\,x$$

Is it true that $\hat{x}_{\text{squared}}$ is the element of Ω which is closest in 2-norm to $\mathrm{E}\,x$?

(e) Let's simulate the random variable x (using your code from another question). Collect $10,000$ samples from x, and compute the error achieved by your estimates $\hat{x}_{\text{min-error}}$, $\hat{x}_{\text{squared}}$ and $\hat{x}_{\text{abs}}$ using the appropriate error measure. Also compute the mean square error achieved by using the mean as an estimate.

2. ***Chernoff Bounds***

(a) The Chernoff bound is similar to the Markov and Chebyshev bounds, as follows. Suppose $x : \Omega \to \mathbb{R}$ is a random variable, and let λ be any positive real number. Then for all $\varepsilon > 0$,

$$\mathrm{Prob}(x \geq \varepsilon) \leq e^{f(\lambda) - \lambda\varepsilon} \qquad (1)$$

where the function f is

$$f(\lambda) = \log \mathrm{E}(e^{\lambda x})$$

Prove this.

(b) Suppose we have n IID discrete random variables $y_1, \ldots, y_n$. Each y_i is a *Bernoulli* random variable, that is it takes value either 0 or 1 with probability $\frac{1}{2}$. Define the random variable x by

$$x = \sum_{i=1}^{n} y_i$$

What is the pmf of x?

(c) When x is as in part (b), what is $f(\lambda)$? Hint; use the fact that x is a sum of IID random variables.

(d) By choosing λ to minimize the right-hand side of equation (1), we can make the bound as tight as possible. Show that the optimal λ is

$$\lambda = \log\left(\frac{\varepsilon}{n-\varepsilon}\right)$$

and hence when $\varepsilon \geq \mathrm{E}\,x$, the best bound possible is

$$\mathrm{Prob}(x \geq \varepsilon) \leq \frac{1}{2^n}\left(\frac{n}{n-\varepsilon}\right)^n \left(\frac{\varepsilon}{n-\varepsilon}\right)^{-\varepsilon}$$

What is the bound when $\varepsilon = \mathrm{E}\,x$?

(e) For $n = 10$, plot

$$\mathrm{Prob}(x \geq \varepsilon)$$

against ε, and also plot the Chernoff bound.

We can also use the Chebyshev bound to give an upper bound on this probability, since if $\varepsilon \geq \mathrm{E}\,x$ we have

$$\begin{aligned} \mathrm{Prob}(x \geq \varepsilon) &= \mathrm{Prob}\big(x - \mathrm{E}\,x \geq \varepsilon - \mathrm{E}\,x\big) \\ &\leq \mathrm{Prob}\big(|x - \mathrm{E}\,x| \geq \varepsilon - \mathrm{E}\,x\big) \\ &\leq \frac{\mathrm{cov}(x)}{(\varepsilon - \mathrm{E}\,x)^2} \end{aligned}$$

Also plot the Chebyshev bound on your plot. Notice that the Chernoff bound is tighter than the Chebyshev bound.