ECSE 305, W09

Assignment #9, Solutions

1. a) The CDF of X and Y are given by

$$F_X(x) = \Phi\left(\frac{x}{\sigma}\right) \qquad (1)$$

$$F_Y(y) \overset{(1)}{=} \begin{cases} 2\Phi\left(\frac{\sqrt{y}}{\sigma}\right) - 1, & y \geq 0 \\ 0, & y < 0 \end{cases} \qquad (2)$$

On the one hand, since $Y = X^2$, we have

$$P(|X| \leq \sigma, Y > \sigma^2) = 0$$

On the other hand, using (1) and (2) above, we have

$$P(|X| \leq \sigma) = 2\Phi(1) - 1 > 0$$

$$P(Y > \sigma^2) = 1 - F_Y(\sigma^2) = 2(1 - \Phi(1)) > 0$$

Since $P(|X| \leq \sigma, Y > \sigma^2) \neq P(|X| \leq \sigma, Y > \sigma^2)$, we conclude that X and Y are not independent

b) We note that $\mu_X = E(X) = 0$ and $\mu_Y = E(X^2) = \sigma^2$. Hence

$$Cov(X,Y) = E(XY) - \cancel{\mu_X}\mu_Y$$

(1) See Section 7.2.1 in the notes.

$$= E(X^3)$$

$$= \int_{-\infty}^{\infty} x^3 \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2} dx \overset{(2)}{=} 0$$

Thus, X and Y are uncorrelated. This example shows that "strongly related" RVs can have zero covariance and be uncorrelated

2. The current is given by $I = V/R$. Since R and V are independent, we have

$$E(I) = E(V/R) = E(V)\,E(1/R)$$

$$E(I^2) = E(V^2/R^2) = E(V^2)\,E(1/R^2)$$

Since $V \sim N(10,1)$, we have $E(V) = 10$ and $E(V^2) = \sigma_V^2 + E(V)^2 = 101$. Since $R \sim U(4,6)$:

$$E(1/R) = \int_4^6 \frac{1}{2r} dr = \frac{1}{2} \ln \frac{3}{2}$$

$$E(1/R^2) = \int_4^6 \frac{1}{2r^2} dr = \frac{1}{24}$$

Finally, $E(I)$

$$E(I) = 5 \ln \frac{3}{2} \cong 2.0273 (A)$$

$$Var(I) = E(I^2) - E(I)^2 \cong 0.0983 (A^2)$$

---

(2) Note that the $N(0,\sigma^2)$ PDF is symmetric, so that all its moments $E(X^k)$, where k = odd integer, are zero.

3. a) $\sum\sum\sum_{\text{all}(x,y,z)} p(x,y,z) = \sum_{x=4}^{5}\sum_{y=1}^{3}\sum_{z=1}^{2} c\,xyz$

$$= c\left(\sum_{x=4}^{5} x\right)\left(\sum_{y=1}^{3} y\right)\left(\sum_{z=1}^{2} z\right)$$

$$= c(9 \times 6 \times 3) = 1$$

$$\Longrightarrow \quad c = \frac{1}{9\times 6\times 3} = \frac{1}{162}$$

b) $p_X(x) = \sum_{y=1}^{3}\sum_{z=1}^{2} c\,xyz$

$$= c\left(\sum_{y=1}^{3} y\right)\left(\sum_{z=1}^{2} z\right) x$$

$$= c\,(6\times 3)\,x = \frac{1}{9}x, \quad x \in R_X$$

$$p_Y(y) = \sum_{x=4}^{5}\sum_{z=1}^{2} c\,xyz = \frac{1}{6}y, \quad y \in R_Y$$

$$p_Z(z) = \sum_{x=4}^{5}\sum_{y=1}^{3} c\,xyz = \frac{1}{3}z, \quad z \in R_Z$$

c) $p_{XY}(x,y) = \sum_{z=1}^{2} c\,xyz = \frac{1}{54}xy, \quad (x,y) \in R_X \times R_Y$

$$p_{YZ}(y,z) = \sum_{x=4}^{5} c\,xyz = \frac{1}{18}yz, \quad (y,z) \in R_Y \times R_Z$$

$$p_{XZ}(x,z) = \sum_{y=1}^{3} c\,xyz = \frac{1}{27}xz, \quad (x,z) \in R_X \times R_Z$$

4. a) $\iiint_{\mathbb{R}^3} f(x,y,z)\,dx\,dy\,dz = \int_0^1\int_0^1\int_0^1 c\,xyz\,dx\,dy\,dz$

$$= c\left(\int_0^1 x\,dx\right)^3$$

$$= c\left(\frac{1}{2}\right)^3 \Longrightarrow c = 8$$

b) The marginal PDF of X is obtained from

$$f_X(x) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y,z)\, dy\, dz$$

If $x \notin (0,1)$, then $f(x,y,z) = 0$ and $f_X(x) = 0$. If $x \in (0,1)$:

$$f_X(x) = \int_0^1 \int_0^1 c\, xyz\, dy\, dz = 8 \left(\int_0^1 y\, dy \right)^2 x = 2x$$

In a similar way, we find

$$f_Y(y) = \begin{cases} 2y, & y \in (0,1) \\ 0, & \text{if not} \end{cases} \qquad f_Z(z) = \begin{cases} 2z, & z \in (0,1) \\ 0, & \text{if not} \end{cases}$$

c) Observe that $f(x,y,z) = f_X(x) f_Y(y) f_Z(z)$ for all $(x,y,z) \in \mathbb{R}^3$. Thus, X, Y and Z are independent and this implies $\rho(X,Y) = \rho(X,Z) = \rho(Y,Z) = 0$.

5. For a randomly selected family, there are 4 possibilities, namely (b,b), (b,g), (g,b) and (g,g) where, for example, (b,g) means that the 1st child (oldest) is a boy and the 2nd is a girl. Because of the independence assumption, each outcome has an equal probability of $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$. We note that

$$P(X_1=1) = P(\{(g,b),(g,g)\}) = \frac{1}{2}$$
$$P(X_2=1) = P(\{(b,g),(g,g)\}) = \frac{1}{2}$$
$$P(X_3=1) = P(\{(b,g),(g,b)\}) = \frac{1}{2}$$

so that $P(X_1=1)\,P(X_2=1)\,P(X_3=1) = 1/8$. However

$$P(X_1=1, X_2=1, X_3=1) = P(\emptyset) = 0$$

Since $P(X_1=1, X_2=1, X_3=1) \neq P(X_1=1)\,P(X_2=1)\,P(X_3=1)$, we conclude that X_1, X_2, X_3 are NOT independent

6. Let X denote the number of letters handled daily. We have $\mu_X = 20000$ and $\sigma_X = 1000$. We seek

$$\begin{aligned} p &\triangleq P(19000 \leq X \leq 21000) \\ &= P(|X-\mu_X| \leq \sigma) \\ &= 1 - P(|X-\mu_X| > \sigma) \end{aligned} \qquad (1)$$

Applying Chebyshev's inequality

$$P(|X-\mu_X| > \sigma) \leq P(|X-\mu_X| \geq \sigma) \leq 1 \qquad (2)$$

Combining (1) and (2), we conclude that $p \geq 0$, which we already know! Therefore, in this case, nothing can be said about p. However, using the same approach, you can show that

$$P(|X-20000| \leq 1000\alpha) \geq 1 - \frac{1}{\alpha^2}$$

which conveys information for any $\alpha > 1$.

7. Let $X_1, \dots, X_n$ be i.i.d., with each X_i uniformly distributed over the interval $(0,1)$. We have (cf. 7)

$$\mu_i = E(X_i) = \frac{1}{2}, \qquad \sigma_i^2 = Var(X_i) = \frac{1}{12}$$

Let $Y = X_1 + X_2 + \dots + X_n$. Since the X_i are independent and therefore uncorrelated (i.e. $Cov(X_i, X_j) = 0$ for $i \neq j$):

$$Var(Y) = Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i) = \frac{n}{12}$$

8. a) $$p(x) = \binom{m}{x} p^x (1-p)^{m-x}, \quad x \in R_X = \{0,1,\dots,m\}$$

$$\Psi(\omega) = \sum_{x=0}^{m} p(x)\, e^{-j\omega x}$$

$$= \sum_{x=0}^{m} \binom{m}{x} p^x q^{m-x} e^{-j\omega x}, \qquad q \triangleq 1-p$$

$$= \sum_{x=0}^{m} \binom{m}{x} (pe^{-j\omega})^x q^{m-x}$$

$$\overset{(2.41)}{=} (pe^{-j\omega} + q)^m \qquad (1)$$

b) Since the X_i are independent, the CF of their sum $Y = X_1 + \dots + X_k$ is obtained as follows:

$$\Psi_Y(\omega) = \Psi_{X_1}(\omega) \dots \Psi_{X_k}(\omega), \quad \text{i.e. product}$$

$$= (pe^{-j\omega}+q)^{m_1} \dots (pe^{-j\omega}+q)^{m_k}$$

$$= (pe^{-j\omega}+q)^{m_1+\dots+m_k} \qquad (2)$$

Comparing (2) to (1), we conclude that $Y \sim B(m', p')$ with parameters $m' = m_1 + m_2 + \dots + m_k$ and $p' = p$.

c) There are $k = 512$ packets. Let X_i denote the number of bits in error in the ith packet ($i = 1, \dots, k$). The X_i are i.i.d. RV, each with distribution $B(m,p)$, $m = 1024$ and $p = 10^{-2}$. The common mean and variance of the X_i are

$$\mu = E(X_i) = mp = 10.24$$
$$\sigma^2 = \mathrm{Var}(X_i) = mpq \simeq 10.14 \qquad (\sigma \simeq 3.2)$$

The total number of bits received in error is $Y = X_1 + X_2 + \dots + X_k$. According to the CLT, Y is approximately $N(k\mu, k\sigma^2)$. Since Y is discrete, we have

$$P(Y < 128) = P(Y \leq 127.5)$$

$$\stackrel{(CLT)}{\cong} \Phi\left(\frac{127.5 - k\mu}{\sqrt{k}\,\sigma}\right)$$

$$\cong \Phi(-70.6) \cong 0$$

9. Each person being polled corresponds to the observation of a RV X_i $(i = 1, \dots, n)$ with Bernoulli distribution, i.e.

$$X_i = \begin{cases} 1, & \text{with prob. } p \\ 0, & \text{" " } q = 1-p \end{cases}$$

Note: $E(X_i) = 1 \times p + 0 \times q = p$
$E(X_i^2) = 1^2 \times p + 0 \times q = p$
$Var(X_i) = E(X_i^2) - E(X_i)^2$
$= p - p^2 = pq$

where p is the true (but unknown) voting intention. The voting intention is estimated as (sample mean)

$$Y = \frac{1}{n}(X_1 + \dots + X_n)$$

If the survey is well-designed, we can assume that the X_i's are independent. Therefore

$$E(Y) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\sum_{i=1}^{n} p = p$$

$$Var(Y) = \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) = \frac{1}{n^2}\sum_{i=1}^{n} pq = \frac{pq}{n}$$

We seek n such that $P(|Y - p| \leq 10^{-2}) \geq \frac{19}{20}$

For n large, we can apply the CLT and Y is approximated as a $N(p, pq/n)$ law. Hence

$$P(|Y-p| \leq 10^{-2}) = P\left(\frac{|Y-p|}{\sqrt{pq/n}} \leq \sqrt{\frac{n}{pq}}\, 10^{-2}\right)$$

$$\overset{CLT}{\simeq} 2\Phi\left(\sqrt{\frac{n}{pq}}\, 10^{-2}\right) - 1 \geq .95$$

$$\Longrightarrow \Phi\left(\sqrt{\frac{n}{pq}}\, 10^{-2}\right) \geq 0.975$$

$$\Longrightarrow \sqrt{\frac{n}{pq}}\, 10^{-2} \geq 1.96 \quad , \text{(from Table, p.196)}$$

$$\Longrightarrow n \geq (196)^2 pq = 38416\, p(1-p)$$

We note that the required value of n depends on p. In the absence of further knowledge, we consider the worst case (i.e. the value of p maximizing $p(1-p)$, which is $p = 1/2$) and so $n \geq 38416/4 = 9604$.

10. a) Let X_i $(i=1,2,\dots)$ be a sequence of uncorrelated RVs with mean $\mu_i = E(X_i)$ and common variance $\sigma^2 = Var(X_i)$. Define $\bar{\mu}_n \triangleq \frac{1}{n}\sum_{i=1}^{n} \mu_i$. For any $\epsilon > 0$, we have

$$\lim_{n\to\infty} P(|\bar{X}_n - \bar{\mu}_n| < \epsilon) = 1 \qquad (1)$$

b) Recall the def. of $\bar{X}_n \triangleq \frac{1}{n}\sum_{i=1}^{n} X_i$. Under the assumption in part a), we have

$$E(\bar{X}_n) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\sum_{i=1}^{n} \mu_i = \bar{\mu}_n$$

$$Var(\bar{X}_n) = \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) = \frac{1}{n^2}\sum_{i=1}^{n} \sigma^2 = \frac{\sigma^2}{n}$$

Applying Chebyshev's inequality, we have for any $\epsilon > 0$:

$$0 \leq P(|\bar{X}_n - \bar{\mu}_n| \geq \epsilon) \leq \frac{\sigma^2}{n\epsilon^2}$$

Taking the limit as $n \to \infty$ on both side:

$$\lim_{n\to\infty} P(|\bar{X}_n - \bar{\mu}_n| \geq \epsilon) = 0 \qquad (2)$$

which is equivalent to (1).