University of Illinois Fall 1998

ECE 313: Solutions to Problem Set #14

1. The easiest way to solve this problem is the indirect way. For each $i = 1, 2, \ldots, n$, we define the random variables:

$$X_i = \begin{cases} 1 & i\text{-th die shows } 1 \\ 0 & \text{otherwise} \end{cases} \qquad Y_i = \begin{cases} 1 & i\text{-th die shows } 2 \\ 0 & \text{otherwise} \end{cases}$$

Since the same die cannot land on both 1 and 2, we conclude that $X_i Y_i = 0$ always. This is a deterministic constant, so that $E[X_i Y_i] = 0$ for all i. It is furthermore easy to see that $E[X_i] = P\{X_i = 1\} = 1/6$ and $E[Y_i] = P\{Y_i = 1\} = 1/6$, for all i. Thus we have:

$$\mathsf{Cov}(X_i, Y_j) = E[X_i Y_j] - E[X_i]E[Y_j] = \begin{cases} -\frac{1}{36} & \text{if } j = i \\ 0 & \text{if } j \neq i \end{cases}$$

The second case follows from the fact that the n rolls of the dice are independent experiments, and therefore X_i and Y_j are independent random variables for $j \neq i$. Now observe that $X = X_1 + X_2 + \cdots + X_n$ and $Y = Y_1 + Y_2 + \cdots + Y_n$. Therefore, we obtain:

$$\mathsf{Cov}(X, Y) = \mathsf{Cov}\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{n} Y_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} \mathsf{Cov}(X_i, Y_j) = \sum_{i=1}^{n} \mathsf{Cov}(X_i, Y_i) = -\frac{n}{36}$$

where the second equality follows from the bilinearity of the covariance function, the third equality follows from the fact that $\mathsf{Cov}(X_i, Y_j) = 0$ unless $j = i$, and the final equality follows from the fact that $\mathsf{Cov}(X_i, Y_i) = -1/36$ for all i.

2. **(a)** Since $Z = 2(X+Y)(X-Y) = 2X^2 - 2Y^2$, we have

$$E[Z] = 2(E[X^2] - E[Y^2]) = 2(\mathsf{Var}(X) - \mathsf{Var}(Y) + E^2[X] + E^2[Y]) = -40$$

(b) We have

$$\begin{aligned} \mathsf{Cov}(T, U) &= \mathsf{Cov}(2X + Y, 2X - Y) \\ &= 4\mathsf{Cov}(X, X) - 2\mathsf{Cov}(X, Y) + 2\mathsf{Cov}(Y, X) - \mathsf{Cov}(Y, Y) \\ &= 4\mathsf{Var}(X) - \mathsf{Var}(Y) = 7 \end{aligned}$$

(c) Here

$$\begin{aligned} E[W] &= E[3X + Y + 2] = 3E[X] + E[Y] + 2 = 9 \\ \mathsf{Var}(W) &= \mathsf{Var}(3X + Y) = 9\mathsf{Var}(X) + \mathsf{Var}(Y) + 6\mathsf{Cov}(X, Y) = 48.6 \end{aligned}$$

(d) If X and Y are jointly Gaussian, then their linear combination $W = 3X + Y + 2$ is a Gaussian random variable. Therefore

$$P\{W > 0\} = 1 - \Phi\left(\frac{0 - E[W]}{\sqrt{\mathsf{Var}(W)}}\right) = 1 - \Phi\left(\frac{0 - 9}{\sqrt{48.6}}\right) = \Phi\left(\frac{9}{\sqrt{48.6}}\right)$$

3. The key point to observe here is that $X_1, X_2, \ldots$ are independent random variables, and therefore Y_n and Y_{n+k} are also independent if $k \geq 3$. Thus $\mathsf{Cov}(Y_n, Y_{n+k}) = 0$ if $k \geq 3$. Now, for $k = 0$, we have:

$$\mathsf{Cov}(Y_n, Y_n) = \mathsf{Var}(Y_n) = \mathsf{Var}(X_n + X_{n+1} + X_{n+2}) = \mathsf{Var}(X_n) + \mathsf{Var}(X_{n+1}) + \mathsf{Var}(X_{n+2}) = 3\sigma^2$$

where we have used the independence of X_n, X_{n+1}, X_{n+2} once again. For $k = 1$, we have:

$$\begin{aligned}
\mathsf{Cov}(Y_n, Y_{n+1}) &= \mathsf{Cov}(X_n + X_{n+1} + X_{n+2}, X_{n+1} + X_{n+2} + X_{n+3}) \\
&= \mathsf{Cov}(X_n, X_{n+1} + X_{n+2} + X_{n+3}) + \mathsf{Cov}(X_{n+1} + X_{n+2}, X_{n+1} + X_{n+2} + X_{n+3}) \\
&= 0 + \mathsf{Cov}(X_{n+1} + X_{n+2}, X_{n+1} + X_{n+2}) + \mathsf{Cov}(X_{n+1} + X_{n+2}, X_{n+3}) \\
&= \mathsf{Var}(X_{n+1} + X_{n+2}) + 0 \;=\; \mathsf{Var}(X_{n+1}) + \mathsf{Var}(X_{n+2}) \;=\; 2\sigma^2
\end{aligned}$$

where we have used the bilinearity of the covariance function, along with the independence of $X_n, X_{n+1}, X_{n+2}, X_{n+3}$. Finally, for $k = 2$, we have:

$$\begin{aligned}
\mathsf{Cov}(Y_n, Y_{n+2}) &= \mathsf{Cov}(X_n + X_{n+1} + X_{n+2}, X_{n+2} + X_{n+3} + X_{n+4}) \\
&= \mathsf{Cov}(X_n + X_{n+1}, X_{n+2} + X_{n+3} + X_{n+4}) + \mathsf{Cov}(X_{n+2}, X_{n+2} + X_{n+3} + X_{n+4}) \\
&= 0 + \mathsf{Cov}(X_{n+2}, X_{n+2}) + \mathsf{Cov}(X_{n+2}, X_{n+3} + X_{n+4}) \\
&= \mathsf{Var}(X_{n+2}) + 0 \;=\; \sigma^2
\end{aligned}$$

where the reasoning is similar to that used for the case $k = 1$. Summarizing all of the above, we conclude that $\mathsf{Cov}(Y_n, Y_{n+k}) = (3 - k)\sigma^2$ for $k = 0, 1, 2$ and $\mathsf{Cov}(Y_n, Y_{n+k}) = 0$ for $k \geq 3$.

4. Define $a = \cos\theta$ and $b = \sin\theta$. Thus a and b are real constants, and we have:

$$\begin{aligned} Z &= aX + bY \\ W &= -bX + aY \end{aligned} \qquad \text{or} \qquad \begin{bmatrix} Z \\ W \end{bmatrix} = \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix}$$

Since this is a linear transformation of jointly Gaussian random variables, we can conclude that Z and W are joinly Gaussian. Therefore, Z and W are independent if and only if they are uncorrelated. Thus, we need to solve

$$\begin{aligned}
\mathsf{Cov}(Z, W) &= \mathsf{Cov}(aX + bY, -bX + aY) \\
&= -ab\mathsf{Cov}(X, X) + a^2\mathsf{Cov}(X, Y) - b^2\mathsf{Cov}(Y, X) + ab\mathsf{Cov}(Y, Y) \\
&= ab\big(\mathsf{Var}(Y) - \mathsf{Var}(X)\big) \;+\; (a^2 - b^2)\mathsf{Cov}(X, Y) \\
&= \cos\theta\,\sin\theta\,(\sigma_2^2 - \sigma_1^2) \;+\; (\cos^2\theta - \sin^2\theta)\,\rho\sigma_1\sigma_2 \\
&= 0.5\sin 2\theta\,(\sigma_2^2 - \sigma_1^2) \;+\; \cos 2\theta\,\rho\sigma_1\sigma_2 \;=\; 0
\end{aligned}$$

for the value of θ. The solution

$$\tan 2\theta \;=\; \frac{2\rho\sigma_1\sigma_2}{\sigma_2^2 - \sigma_1^2}$$

specifies the desired angle θ. Note that for $\sigma_1 = \sigma_2$, we get $\theta = \pi/4$.

5. **(a)** We have

$$\mu \;=\; E[X] \;=\; \int_2^\infty u 24u^{-4}\,du \;=\; 24\int_2^\infty u^{-3}\,du \;=\; 3$$

Thus

$$\begin{aligned}
q(\delta) &= P\{|X - 3| > \delta\} \;=\; 1 - P\{|X - 3| < \delta\} \\
&= 1 - \int_{\min(2, 3-\delta)}^{3+\delta} 24u^{-4}\,du \;=\; \begin{cases} 1 - \frac{8}{(3-\delta)^3} + \frac{8}{(3+\delta)^3} & \delta < 1 \\ \frac{8}{(3+\delta)^3} & \delta \geq 1 \end{cases}
\end{aligned}$$

(b) Using Chebyshev's inequality, we obtain

$$q(\delta) \;=\; P\{|X-\mu| > \delta\} \;\leq\; \frac{\mathrm{Var}(X)}{\delta^2} \;=\; \frac{3}{\delta^2}$$

where we have used the fact that $\mathrm{Var}(X) = E[X] = 3$ for this density.

(c) Plot not provided. $\frac{3}{\delta^2}$ will typically be much larger than $q(\delta)$, as you can see.

6. Extra Credit: By the weak law of large numbers,

$$P\left\{\, |\widehat{p} - p| \geq 0.02 \,\right\} \;\leq\; \frac{\mathrm{Var}(X_i)}{n(0.02)^2}$$

The variance of a Bernoulli random variable with mean p is given by $\mathrm{Var}(X_i) = p - p^2$. Thus, to guarantee the $\pm 2\%$ margin of error with probability $1 - 0.05 = 0.95$, we need

$$\frac{p - p^2}{n(0.02)^2} \;\leq\; 0.05$$

or, equivalently, $n \geq\; 50000(p - p^2)$. For $p = 0.3$ this reduces to $N = 50000 \cdot 0.21 = 10500$. In general, if p is unknown, the poll has to be designed for that value of p which maximizes the function $p - p^2$. Since $p - p^2$ is a simple quadratic in p, it is easy to see that this value is $p = 0.5$. Thus, the Sirrah Poll will have to survey $50000 \cdot (0.5 - 0.25) = 12500$ voters.