

1. Correlation of histogram values

- (a) X_1 is Bernoulli with parameter $p = \frac{1}{6}$, so $E[X_1] = \frac{1}{6}$ and $Var(X_1) = \frac{1}{6}(1 - \frac{1}{6}) = \frac{5}{36}$.
 (b) $E[X] = nE[X_1] = \frac{n}{6}$ and $Var(X) = nVar(X_1) = \frac{5n}{36}$.
 (c) We begin by computing $Cov(X_1, Y_1)$. Since $X_1Y_1 = 0$ with probability one, $E[X_1Y_1] = 0$. Therefore $Cov(X_1, Y_1) = E[X_1Y_1] - E[X_1]E[Y_1] = 0 - \frac{1}{6}\frac{1}{6} = -\frac{1}{36}$. So $Cov(X_i, Y_i) = -\frac{1}{36}$ for any i . On the other hand, if $i \neq j$ then X_i is independent of X_j . So

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

(d)

$$Cov(X, Y) = \sum_i \sum_j Cov(X_i, Y_j) = \sum_i Cov(X_i, Y_i) = nCov(X_1, Y_1) = -\frac{n}{36}$$

and

$$\rho(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var(X)Var(Y)}} = \frac{-1}{5}$$

2. Conditional expectation of histogram values

- (a) Given that x of the dice show a 1, each of the remaining dice is equally likely to show 2, 3, 4, 5, or 6. Thus, each of the remaining $n - x$ dice shows a 2 with conditional probability $\frac{1}{5}$. Therefore $E[Y|X = x] = \frac{n-x}{5}$.
 (b) This is the same function found in part (a) by general theory. That is, $g(x) = E[Y|X = x] = \frac{n-x}{5}$.

3. Covariance for a Poisson process

- (a) As found in Problem 3(d) of Problem Set 10, the conditional distribution of $N(\tau)$ given $N(T) = n$ is binomial with parameters n and $p = \frac{\tau}{T}$. Therefore $E[N(\tau)|N(T) = n] = \frac{n\tau}{T}$.
 (b) As found in Problem 3(e) of Problem Set 10, the conditional distribution of $N(T)$ given $N(\tau) = i$ is the same as the distribution of i plus a Poisson random variable with mean $(T - \tau)\lambda$. Therefore $E[N(T)|N(\tau) = i] = i + (T - \tau)\lambda$.
 (c) Since $N(\tau)$ is a Poisson random variable with mean $\lambda\tau$, and since the variance of a Poisson random variable is the same as the mean, $Cov(N(\tau), N(\tau)) = Var(N(\tau)) = \lambda\tau$.
 (d) Since $N(\tau)$ is independent of $N(T) - N(\tau)$, $Cov(N(\tau), N(T) - N(\tau)) = 0$.
 (e) Use $N(T) = N(\tau) + (N(T) - N(\tau))$ to get $Cov(N(\tau), N(T)) = Cov(N(\tau), N(\tau) + (N(T) - N(\tau))) = Cov(N(\tau), N(\tau)) + Cov(N(\tau), (N(T) - N(\tau))) = \lambda\tau + 0 = \lambda\tau$.

4. Conditional expectation for uniform density over a triangle

- (a) The triangle has base and height one, so the area of the triangle is 0.5. Thus the joint pdf is 2 inside the triangle.
 (b)

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y) dy = \begin{cases} \int_0^{x/2} 2 dy = x & \text{if } 0 < x < 1 \\ \int_{x-1}^{x/2} 2 dy = 2 - x & \text{if } 1 < x < 2 \\ 0 & \text{else} \end{cases}$$

- (c) In view of part (c), the conditional density $f_{Y|X}(y|x)$ is not well defined unless $0 < x < 2$. In

general we have

$$f_{Y|X}(y|x) = \begin{cases} \frac{2}{x} & \text{if } 0 < x \leq 1 \text{ and } y \in [0, \frac{x}{2}] \\ 0 & \text{if } 0 < x \leq 1 \text{ and } y \notin [0, \frac{x}{2}] \\ \frac{2}{2-x} & \text{if } 1 < x < 2 \text{ and } y \in [x-1, \frac{x}{2}] \\ 0 & \text{if } 1 < x < 2 \text{ and } y \notin [x-1, \frac{x}{2}] \\ \text{not defined} & \text{if } x \leq 0 \text{ or } x \geq 2 \end{cases}$$

Thus, for $0 < x \leq 1$, the conditional distribution of Y is uniform over the interval $[0, \frac{x}{2}]$. For $1 < x \leq 2$, the conditional distribution of Y is uniform over the interval $[x-1, \frac{x}{2}]$.

(d) Finding the midpoints of the intervals that Y is conditionally uniformly distributed over, or integrating x against the conditional density found in part (c), yields:

$$E[Y|X = x] = \begin{cases} \frac{x}{4} & \text{if } 0 < x \leq 1 \\ \frac{3x-2}{4} & \text{if } 1 < x < 2 \\ \text{not defined} & \text{if } x \leq 0 \text{ or } x \geq 2 \end{cases}$$

5. Rock and roll

(a) N is the number of independent trials until a success, and the probability of a success for each trial is $\frac{1}{6}$. Thus, N has the geometric distribution with parameter $p = \frac{1}{6}$. Thus, $p_N(n) = \frac{5}{6}^{n-1} \frac{1}{6}$ for $n = 1, 2, 3, \dots$ and $p_N(n) = 0$ for other values of n . As seen on Ross, page 159: $E[N] = 6$, and $Var(N) = \frac{1-p}{p^2} = 30$.

(b) Given $N = n$, S is one more than the sum of $n-1$ random variables $Y_1, \dots, Y_{n-1}$ that are independent and take the values 2, 3, 4, 5, or 6 each with probability $\frac{1}{5}$. The conditional mean of each of these $n-1$ variables is 4, so that $E[S|N = n] = 1 + 4(n-1)$.

(c) $E[S] = E[E[S|N]] = E[1 + 4(N-1)] = 1 + 4E[N] - 4 = 21$. Imagine rolling a die repeatedly without stopping. By the law of large numbers, the frequency of each of the six numbers is the same. Thus, on average, between each two appearances of a 1, any other given number should appear an average of one time.

6. Normal approximation for quantization error

The mean of each roundoff error is zero and the variance is $\int_{-0.5}^{0.5} u^2 du = \frac{1}{12}$. Thus, $E[S] = 0$ and $Var(S) = \frac{100}{12} = 8.333$. Thus, $P[|S| \geq 5] = P[|\frac{S}{\sqrt{8.333}}| \geq \frac{5}{\sqrt{8.333}}] \approx 2Q(\frac{5}{\sqrt{8.333}}) = 2Q(1.73) = 2(1 - \Phi(1.732)) = 0.083$.

7. A comparison of bounds and the normal approximation

(a) $E[X_1] = \frac{1+2+3+4+5+6}{6} = 3.5$ so by Markov's inequality, $P[S_{100} \geq 400] \leq \frac{100(3.5)}{400} = \frac{7}{8}$. (Not at all a strong bound.)

(b) $E[X_1^2] = \frac{1^2+2^2+3^2+4^2+5^2+6^2}{6} = 15$ so that $Var(X_i) = 15 - 3.5^2 = 2.75$ Therefore by Chebyshev's inequality $P[S_{100} \geq 400] = P[S_{100} - 350 \geq 50] \leq P[|S_{100} - 350| \geq 50] \leq \frac{100(2.75)}{50^2} = 0.11$

Or, we could use the one-side Chebyshev inequality found on page 418 of Ross. That yields $P[S_{100} \geq 400] = P[S_{100} - 350 \geq 50] \leq \frac{275}{275+2500} = 0.0991$. (Either answer is acceptable.) (c) The normal approximation based on the central limit theorem yields

$$P[S_{100} \geq 400] = P[\frac{S_{100} - 350}{\sqrt{(2.75)100}} \geq \frac{400 - 350}{\sqrt{(2.75)100}}] \approx Q(\frac{400 - 350}{\sqrt{(2.75)100}}) = Q(3.015) = 1 - \Phi(3.015) \approx 0.0009$$