1. Suppose the number N of times a fair die is rolled is Poisson(λ). Let Y be the total score in N rolls.

We know that the die score X has

$$\mu = \mathbf{E}(X) = \frac{7}{2}, \sigma^2 = \mathrm{Var}(X) = \frac{35}{12}, G_X(s) = \frac{1}{6}\sum_{k=1}^{6} s^k$$

(a) Find $\mathbf{E}(Y|N=n)$, $\mathbf{E}(Y|N)$ and $\mathbf{E}(Y)$. (b) Find $\mathbf{E}\left(Y^2|N=n\right)$, $\mathbf{E}\left(Y^2|N\right)$ and $\mathbf{E}\left(Y^2\right)$. Find Var($Y$) .

Answer. **Given** $N = n$**,** we have $Y = X_1 + \ldots X_n$, where X_i are independent die scores (distributed like X). Hence,

$$\begin{aligned} \mathbf{E}(Y|N=n) &= \mathbf{E}(X_1) + \ldots + \mathbf{E}(X_n) = n\mu, \\ \mathbf{E}\left(Y^2|N=n\right) &= \mathrm{Var}(X_1 + \ldots + X_n) + (n\mu)^2 = n\sigma^2 + n^2\mu^2. \end{aligned}$$

Hence

$$\mathbf{E}(Y|N) = \mu N, \mathbf{E}(Y) = \mu\mathbf{E}(N) = \lambda\mu,$$

and

$$\begin{aligned} \mathbf{E}\left(Y^2|N\right) &= \sigma^2 N + \mu^2 N^2 \\ \mathbf{E}\left(Y^2\right) &= \sigma^2\mathbf{E}(N) + \mu^2\mathbf{E}\left(N^2\right) = \lambda\sigma^2 + \mu^2\left(\lambda + \lambda^2\right). \end{aligned}$$

Finally, $\mathrm{Var}(Y) = \lambda\sigma^2 + \mu^2\left(\lambda + \lambda^2\right) - (\lambda\mu)^2 = \lambda\left(\sigma^2 + \mu^2\right)$.

(c) Find the generating function of Y.

Answer. By definition, denoting X_k the score in the kth roll,

$$\begin{aligned} G_Y(s) &= \mathbf{E}\left(s^Y\right) = \mathbf{E}\left[\mathbf{E}\left(s^Y|N\right)\right] = \sum_{n=0}^{\infty}\mathbf{E}\left[\mathbf{E}\left(s^Y|N=n\right)\right]\mathbf{P}(N=n) \\ &= \sum_{n=0}^{\infty}\mathbf{E}\left[\mathbf{E}\left(s^{\sum_{k=1}^{n}X_k}\right)\right]\mathbf{P}(N=n) = \sum_{n=0}^{\infty}\left(G_X(s)^n\right)\mathbf{P}(N=n) \\ &= G_N(G_X(s)) = \exp\left\{\lambda\left(\frac{1}{6}\sum_{k=1}^{6}s^k - 1\right)\right\}, \end{aligned}$$

because

$$G_N(s) = \exp\{\lambda(s-1)\}.$$

2. A coin-making machine produces quarters in such way that, for each coin, the probability U to turn up heads is uniform in $(0, 1)$. A coin pops out (randomly) and is tossed multiple times.

(a) Compute the probability that the first two tosses are both heads. Let X_n be the number of heads in the first n tosses. Compute $\mathbf{P}(X_n = k)$ for all $0 \leq k \leq n$.

(b) Let N be the number of tosses needed to get heads for the first time. Compute $\mathbf{P}(N = n)$ for all $n \geq 1$. Compute the expected value of N.

HINT: for all nonnegative integers m, l,

$$\int_0^1 x^m (1-x)^l \, dx = \frac{m!l!}{(m+l+1)!}.$$

Answer. (a) First,

$$\mathbf{P}(H_1 H_2) = \mathbf{E}[\mathbf{P}(H_1 H_2 | U)] = \mathbf{E}\left(U^2\right) = \int_0^1 u^2 du = \frac{1}{3}.$$

Then, for $k = 0, \ldots, n$,

$$\begin{aligned} \mathbf{P}(X_n = k) &= \mathbf{E}[\mathbf{P}(X_n = k | U)] = \mathbf{E}\left[\binom{n}{k} U^k (1-U)^{n-k}\right] \\ &= \binom{n}{k} \int_0^1 u^k (1-u)^{n-k} \, du = \frac{n!}{(n-k)!k!} \frac{k!(n-k)!}{(n+1)!} = \frac{1}{n+1}. \end{aligned}$$

(b) By conditioning with respect to U, for $n \geq 1$,

$$\begin{aligned} \mathbf{P}(N = n) &= \mathbf{E}[\mathbf{P}(N = n | U)] = \int_0^1 \mathbf{P}(N = n | U = u) \, du \\ &= \int_0^1 u^{n-1} (1-u) = \frac{(n-1)!1!}{(n+1)!} = \frac{1}{n(n+1)}, \end{aligned}$$

and

$$\mathbf{E}(N) = \sum_{n=1}^{\infty} n \frac{1}{n(n+1)} = \sum_{n=1}^{\infty} \frac{1}{n+1} = \infty.$$

3. An urn contains $2n$ balls, coming in pairs: two balls are labeled "1", two balls are labeled "2",..., two balls are labeled "n", $n > 7$. A sample of size n is taken without replacement.

a) Denote by N the number of pairs in the sample. Compute the expected value and the variance of N. You do not need to simplify the expression for the variance.

Answer. Let A_i ="ith pair selected", $i = 1, \ldots, n$. Then $N = \sum_{i=1}^{n} I_{A_i}$, and

$$
\begin{aligned}
\mathbf{E}(N) &= n\mathbf{P}(A_1) = \frac{n(n-1)}{2(2n-1)}, \\
\operatorname{Var}(N) &= \sum_{i=1}^{n} \operatorname{Var}\left(I_{A_i}\right) + 2\sum_{i<j} \operatorname{Cov}\left(I_{A_i}, I_{A_j}\right) = n\operatorname{Var}(I_{A_1}) + 2\binom{n}{2}\operatorname{Cov}(I_{A_1}, I_{A_2}),
\end{aligned}
$$

where $\operatorname{Var}(I_{A_1}) = \mathbf{P}(A_1) - \mathbf{P}(A_1)^2$,$\operatorname{Cov}(I_{A_1}, I_{A_2}) = \mathbf{P}(A_1A_2) - \mathbf{P}(A_1)^2$.

We find

$$\mathbf{P}(A_1) = \frac{\binom{2n-2}{n-2}}{\binom{2n}{n}} = \frac{n-1}{2(2n-1)}$$

and

$$\mathbf{P}(A_1 \cap A_2) = \frac{\binom{2n-4}{n-4}}{\binom{2n}{n}} = \frac{(n-2)(n-3)}{2^2(2n-1)(2n-3)}.$$

Similarly,

$$\mathbf{P}(A_1A_2A_3) = \frac{\binom{2n-6}{n-6}}{\binom{2n}{n}} = \frac{(n-3)(n-4)(n-5)}{2^3(2n-1)(2n-3)(2n-5)}.$$

b) Find probability that at least one of the first three pairs is selected.

Answer. By Inclusion/exclusion principle, with probabilities found above,

$$
\begin{aligned}
\mathbf{P}(A_1 \cup A_2 \cup A_3) &= \mathbf{P}(A_1) + \mathbf{P}(A_2) + \mathbf{P}(A_3) - \mathbf{P}(A_1A_2) - \mathbf{P}(A_1A_3) - \mathbf{P}(A_2A_3) \\
+\mathbf{P}(A_1A_2A_3) &= 3\mathbf{P}(A_1) - 3\mathbf{P}(A_1A_2) + \mathbf{P}(A_1A_2A_3)
\end{aligned}
$$

4. Let $Y = X + \varepsilon Z$, where $\varepsilon > 0$, and $X \sim N\left(\mu, \sigma^2\right), Z \sim N(0,1)$ are independent.

(a) Find $\operatorname{Var}(Y)$,$\operatorname{Cov}(X, Y)$ and the correlation coefficient $\rho = \rho(X, Y)$;

Answer. Because of independence,

$$
\begin{aligned}
\operatorname{Var}(Y) &= \sigma^2 + \varepsilon^2, \operatorname{Cov}(X, Y) = \operatorname{Cov}(X, X + \varepsilon Z) = \sigma^2, \\
\rho &= \frac{\operatorname{Cov}(X, Y)}{\sigma\sqrt{\sigma^2 + \varepsilon^2}} = \frac{\sigma}{\sqrt{\sigma^2 + \varepsilon^2}}.
\end{aligned}
$$

(b) Is (X, Y) normal bivariate? Find $\mathbf{E}(X|Y)$ and the error $\mathbf{E}\left[\left(\hat{X}-X\right)^2\right]$ with $\hat{X}=\mathbf{E}(X|Y)$.

Answer. (X, Y) is normal bivariate Since (X, Z) is normal bivariate as a pair of independent normal, and

$$(X, Y) = (X, Z)\begin{pmatrix} 1 & 1 \\ 0 & \varepsilon \end{pmatrix}$$

with an invertible matrix. We know its parameters, $\mathbf{E}(Y) = \mathbf{E}(X) = \mu$: given $Y = y$, we have $X \sim N\left(\rho\frac{\sigma}{\sqrt{\sigma^2+\varepsilon^2}}(y-\mu)+\mu, \sigma^2\left(1-\rho^2\right)\right)$. Hence

$$\begin{aligned}
\hat{X} &= \mathbf{E}(X|Y) = \rho\frac{\sigma}{\sqrt{\sigma^2+\varepsilon^2}}(Y-\mu)+\mu, \\
\mathrm{Var}(X|Y) &= \sigma^2\left(1-\rho^2\right).
\end{aligned}$$

Then the error

$$\begin{aligned}
X-\hat{X} &= X-\frac{\sigma^2}{\sigma^2+\varepsilon^2}(X+\varepsilon Z-\mu)-\mu = \left(1-\frac{\sigma^2}{\sigma^2+\varepsilon^2}\right)X-\frac{\sigma^2}{\sigma^2+\varepsilon^2}\varepsilon Z \\
&\quad +\mu\left(\frac{\sigma^2}{\sigma^2+\varepsilon^2}-1\right),
\end{aligned}$$

and the mean square error is (using independence of X, Z),

$$\begin{aligned}
\mathbf{E}\left[\left(X-\hat{X}\right)^2\right] &= \mathrm{Var}\left(X-\hat{X}\right) = \left(1-\frac{\sigma^2}{\sigma^2+\varepsilon^2}\right)^2\sigma^2+\left(\frac{\sigma^2}{\sigma^2+\varepsilon^2}\right)^2\varepsilon^2 \\
&= \frac{\varepsilon^4\sigma^2}{\left(\sigma^2+\varepsilon^2\right)^2}+\frac{\sigma^4\varepsilon^2}{\left(\sigma^2+\varepsilon^2\right)^2} = \varepsilon^2\frac{\sigma^2}{\sigma^2+\varepsilon^2}.
\end{aligned}$$

Alternatively, $\mathbf{E}\left[\left(X-\hat{X}\right)^2\right] = \mathbf{E}\{\mathrm{Var}(X|Y)\} = \sigma^2\left(1-\rho^2\right) = \varepsilon^2\frac{\sigma^2}{\sigma^2+\varepsilon^2}$.

Comment. If Y itself is used, then the error

$$\mathbf{E}\left[(Y-X)^2\right] = \varepsilon^2 > \varepsilon^2\frac{\sigma^2}{\sigma^2+\varepsilon^2}.$$

5. a) Let X, Y be independent uniform in $(0, 1)$. What is the set of possible values of $V = X + Y$? Find the pdf of V.

Answer. The range of V is $(0,2)$. If f is the pdf of the uniform in $(0,1)$, then

$$\begin{aligned} f_V(x) &= \int_{-\infty}^{\infty} f(x-y) f(y)\, dy = \int_0^1 f(x-y)\, dy \\ &= \int_0^1 I_{(0,1)}(x-y)\, dy = \int_0^1 I_{\{y:\, x-1<y<x\}} dy, x \in \mathbf{R}. \end{aligned}$$

Hence

$$f_V(x) = \begin{cases} \int_0^x dy = x, & x \in (0,1), \\ \int_{x-1}^1 dy = 2-x, & x \in (1,2) \end{cases}$$

b) A stick of length 1 is broken at a point X uniformly distributed over its length. We can assume X is uniform in $(0,1)$ and the length of the longer piece is $L = \max\{X, 1-X\}$. Find the df and pdf of L, $\mathbf{E}(L)$ and Var(L). What is the expected length of a shorter piece?

Answer. The range of L is $(1/2, 1)$. For $x \in (1/2, 1)$,

$$\begin{aligned} F_L(x) &= \mathbf{P}(L \le x) = \mathbf{P}(\max\{X, 1-X\} \le x) = \mathbf{P}(X \le x, 1-X \le x) \\ &= \mathbf{P}(1-x \le X \le x) = \frac{x-(1-x)}{1} = 2x-1, \end{aligned}$$

and $F_L(x) = 0$ if $x < 1/2$, $F_L(x) = 1$ if $x \ge 1$. The pdf

$$f_L(x) = F'(x) = 2, x \in (1/2, 1).$$

That is L is uniform in $(1/2, 1)$. Hence $\mathbf{E}(L) = 3/4$,Var$(L) = \frac{(1-1/2)^2}{12} = \frac{1}{48}$. Expected length of the shorter piece is $1 - 3/4 = 1/4$.

6. a) Let X be Poisson(λ). Show that $\frac{X}{\lambda} \xrightarrow{D} 1$ as $\lambda \to \infty$.

Answer. Recall the characteristic function of a constant c is $\phi_c(t) = e^{ict}$. Since X is Poisson(λ),

$$\phi_{X/\lambda}(t) = \phi_X(t/\lambda) = \exp\left\{\lambda\left(e^{\frac{it}{\lambda}} - 1\right)\right\}.$$

Now

$$\lambda\left(e^{\frac{it}{\lambda}} - 1\right) = \frac{e^{\frac{it}{\lambda}} - 1}{\frac{it}{\lambda}} it \to it$$

as $\lambda \to \infty$. So,

$$\phi_{X/\lambda}(t) = \exp\left\{\lambda\left(e^{\frac{it}{\lambda}} - 1\right)\right\} \to e^{it} = \phi_1(t), t \in \mathbf{R}.$$

Thus $\frac{X}{\lambda} \xrightarrow{D} 1$ as $\lambda \to \infty$ by continuity theorem.

b) Let Y_n be number of "1" in n rolls of a fair die. Approximate for large n the probability

$$\mathbf{P}\left(0.15 < \frac{Y_n}{n} \le 0.2\right)$$

using the distribution function $\Phi(x)$ of a standard normal r.v. Is this probability larger with $n = 100$ or $n = 1000$?

Answer. Y_n is Binomial(n, p) with $p = 1/6$. By CLT for binomial r.v., for large n

$$\begin{aligned} & \mathbf{P}\left(0.15 < \frac{Y_n}{n} \le 0.2\right) \\ \approx\ & \mathbf{P}\left(\frac{0.1 - p}{\sqrt{\frac{p(1-p)}{n}}} \le Z \le \frac{0.2 - p}{\sqrt{\frac{p(1-p)}{n}}}\right) \\ =\ & \Phi\left(\sqrt{n}\frac{0.1 - p}{\sqrt{p(1-p)}}\right) - \Phi\left(\sqrt{n}\frac{0.1 - p}{\sqrt{p(1-p)}}\right) \end{aligned}$$

The larger n, the larger interval for Z : this probability larger with $n = 1000$.

c) Let $X_1, X_2, \ldots$ be independent identically distributed with characteristic function φ. Let N be independent of X_i's with $\mathbf{P}(N = n) = 2^{-n}, n \ge 1$. Let $V = \sum_{i=1}^{N} X_i$. Find the characteristic function of V.

Answer. Given $N = n$, we have $V = \sum_{i=1}^{n} X_i$ in distribution with i.i.d. X_i's, and

$$\begin{aligned} \mathbf{E}\left(e^{itV} | N = n\right) &= \mathbf{E}\left(e^{it\sum_{i=1}^{n} X_i}\right) = \left\{\mathbf{E}\left(e^{itX_1}\right)\right\}^n = \varphi(t)^n, \\ \mathbf{E}\left(e^{itV} | N\right) &= \varphi(t)^N \end{aligned}$$

So, using conditioning,

$$\begin{aligned} \phi_V(t) &= \mathbf{E}\left(e^{itV}\right) = \mathbf{E}\left[\mathbf{E}\left(e^{itV}|N\right)\right] = \mathbf{E}\left(\varphi(t)^N\right) \\ &= \sum_{n=1}^{\infty} \varphi(t)^n \mathbf{P}(N = n) = \sum_{n=1}^{\infty} \varphi(t)^n 2^{-n} = \sum_{n=1}^{\infty} \left(\frac{\varphi(t)}{2}\right)^n \\ &= \frac{\varphi(t)/2}{1 - \varphi(t)/2} = \frac{\varphi(t)}{2 - \varphi(t)}. \end{aligned}$$