1. A box contains 5 fair coins and other 4 coins with heads probability 1/3.

a) One coin is randomly selected and tossed twice. We see heads in the first toss and tails in the second toss. Find probability that the coin is fair.

Answer. Consider the events: $H_1 =$ "heads in the 1st toss", $T_2 =$ "tails in the 2nd toss", $N =$ "selected coin is fair", $N^c =$ "selected coin is not fair". By Bayes formula

$$\begin{aligned}\mathbf{P}(N|H_1T_2) &= \frac{\mathbf{P}(H_1T_2|N)\mathbf{P}(N)}{\mathbf{P}(H_1T_2|N)\mathbf{P}(N)+\mathbf{P}(H_1T_2|N^c)\mathbf{P}(N^c)}\\ &= \frac{\left(\frac{1}{2}\right)^2\frac{5}{9}}{\left(\frac{1}{2}\right)^2\frac{5}{9}+\frac{1}{3}\cdot\frac{2}{3}\cdot\frac{4}{9}} = \frac{45}{77}.\end{aligned}$$

b) We throw away this coin and choose randomly another coin from the 8 coins in the box and toss it. What is the probability to see heads?

Answer. Consider $H_3 =$ "heads in the toss of another coin coin". Then by total probability law 3 times,

$$\begin{aligned}\mathbf{P}(H_3|H_1T_2) &= \mathbf{P}(H_3|N)\mathbf{P}(N|H_1T_2)+\mathbf{P}(H_3|N^c)\mathbf{P}(N^c|H_1T_2)\\ &= \left(\frac{4}{8}\frac{1}{2}+\frac{4}{8}\frac{1}{3}\right)\cdot\frac{45}{77}+\left(\frac{5}{8}\frac{1}{2}+\frac{3}{8}\frac{1}{3}\right)\left(1-\frac{45}{77}\right)=\frac{131}{308}.\end{aligned}$$

2. The amount X of money a customer spends in a certain store is $N\left(\mu,\sigma^2\right)$. Number N of customer arrivals per day to that store is Poisson(λ).

a) Find the moment generating function for the amount Y of money spent daily by customers. Hint: $Y=\sum_{i=1}^N X_i$.

Answer. Given $N=n$, the total amount spent by customers $Y=\sum_{i=1}^n X_i$, where X_i is the amount spent by customer i. Hence

$$\mathbf{E}\left(e^{tY}|N=n\right)=M_X(t)^n, \mathbf{E}\left(e^{tY}|N\right)=M_X(t)^N,$$

and

$$M_Y(t)=\mathbf{E}\left(e^{tY}\right)=\mathbf{E}\left[\mathbf{E}\left(e^{tY}|N\right)\right]=\mathbf{E}\left(M_X(t)^N\right)=G_N(M_X(t)),$$

where (because $X\sim N\left(\mu,\sigma^2\right)$),

$$M_X(t)=e^{\mu t+\sigma^2t^2/2}, t\in\mathbf{R}, G_N(s)=\exp\{\lambda(s-1)\}, s\in\mathbf{R}.$$

Thus

$$M_Y(t)=\exp\left\{\lambda\left(e^{\mu t+\sigma^2t^2/2}\right)-1\right\}, t\in\mathbf{R}.$$

b) Find $\mathbf{E}(Y|N), \mathbf{E}(Y^2|N)$.

Answer. Since given $N = n$, we have $Y = \sum_{i=1}^{n} X_i$, and X_i are independent,

$$\begin{aligned}
\mathbf{E}(Y|N=n) &= \sum_{i=1}^{n} \mathbf{E}(X_i) = n\mathbf{E}(X) = n\mu, \mathbf{E}(Y|N) = \mu N, \\
\mathrm{Var}(Y|N=n) &= \sum_{i=1}^{n} \mathrm{Var}(X_i) = n\sigma^2, \mathrm{Var}(Y|N) = \sigma^2 N, \\
\mathbf{E}(Y^2|N) &= \mathrm{Var}(Y|N) + (\mathbf{E}(Y|N))^2 = \sigma^2 N + N^2\mu^2.
\end{aligned}$$

c) Find $\mathbf{E}(Y)$ and $\mathrm{Var}(Y)$.

Answer. By b),

$$\begin{aligned}
\mathbf{E}(Y) &= \mathbf{E}[\mathbf{E}(Y|N)] = \mathbf{E}(\mu N) = \mu\mathbf{E}(N) = \mu\lambda, \\
\mathrm{Var}(Y) &= \mathbf{E}[\mathrm{Var}(Y|N)] + \mathrm{Var}(\mathbf{E}(Y|N)) = \sigma^2\lambda + \mu^2\lambda.
\end{aligned}$$

`Alternatively,` see the class note of 11/12, Ex.1.

3. Suppose that m balls are placed at random into n boxes (with all n^m possibilities equally likely).

a) Find $\mathbf{P}$(box 1 is not empty, box 2 is not empty, box 3 is not empty) and $\mathbf{P}$(box 1 contains exactly 3 balls).

Answer. Let B_i = "ith box is not empty". Then

$$\mathbf{P}(B_1B_2B_3) = 1 - \mathbf{P}(B_1^c \cup B_2^c \cup B_3^c),$$

where, by inclusion/exclusion and independence,

$$\begin{aligned}
\mathbf{P}(B_1^c \cup B_2^c \cup B_3^c) &= 3\mathbf{P}(B_1^c) - 3\mathbf{P}(B_1^c B_2^c) + \mathbf{P}(B_1^c B_2^c B_3^c) \\
&= 3\left(1 - \frac{1}{n}\right)^m - 3\left(1 - \frac{2}{n}\right)^m + \left(1 - \frac{3}{n}\right)^m.
\end{aligned}$$

The number of balls in box 1 is binomial$(m, p = 1/n)$:

$$\mathbf{P}(\text{box 1 contains exactly 3 balls}) = \binom{m}{3}\left(\frac{1}{n}\right)^3\left(1 - \frac{1}{n}\right)^{m-3}.$$

b) Let X be the number of empty boxes. Compute $\mathbf{E}(X)$ and $\mathrm{Var}(X)$.

Answer. Let A_i = "ith box is empty. We found in part a),

$$p := \mathbf{P}(A_1) = \left(1 - \frac{1}{n}\right)^m, r := \mathbf{P}(A_1A_2) = \left(1 - \frac{2}{n}\right)^m.$$

Since

$$X = \sum_{i=1}^{n} I_{A_i},$$

we find

$$\mathbf{E}(X) = \sum_{i=1}^{n} \mathbf{E}(I_{A_i}) = \sum_{i=1}^{n} \mathbf{P}(A_i) = n\mathbf{P}(A_1) = np.$$

Since

$$X^2 = \sum_{i=1}^{n} I_{A_i}^2 + 2\sum_{i<j} I_{A_i} I_{A_j} = \sum_{i=1}^{n} I_{A_i} + 2\sum_{i<j} I_{A_i} I_{A_j},$$

we have

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

4. Let X, Y be independent identically distributed exponential random variables with parameter $\lambda = 1$.

a) Find the joint probability density function of $\left(\frac{X}{X+Y}, X+Y\right)$.

Answer. Let $U = X/(X+Y), V = X+Y$. We look at (U, V) as a function of (X, Y). First, the joint pdf of (X, Y) is

$$f(x, y) = e^{-x} e^{-y} = e^{-(x+y)}, (x, y) \in D = \{(x, y) : x > 0, y > 0\},$$

and (U, V) is the function of (X, Y) defined by

$$u = \frac{x}{x+y}, v = x+y, x, y > 0. \tag{1}$$

We find its inverse by solving (1) for x, y :

$$x = uv, y = v - x = v - uv \tag{2}$$

and noticing that according to (2), $D = \{(x, y) : x > 0, y > 0\}$ is mapped onto

$$S = \{(u, v) : x = uv > 0, y = v - uv > 0\} = \{(u, v) : 0 < u < 1, v > 0\}.$$

The Jacobian of the inverse

$$J(u, v) = \begin{vmatrix} v & u \\ -v & 1-u \end{vmatrix} = v - vu + vu = v.$$

By Thm in the class note of 11/8 (see Thm in 4.7 of textbook), the joint pdf of (U,V) is

$$\begin{aligned} g(u,v) &= f(uv, v-vu)\,|J(u,v)|\, I_S(u,v) = e^{-v} v I_{\{0<u<1, v>0\}} = I_{\{0<u<1\}} e^{-v} v I_{\{v>0\}} \\ &= ve^{-v}, v>0, 0<u<1. \end{aligned}$$

b) Find and identify the marginal probability density functions. Are $\frac{X}{X+Y}$ and $X+Y$ independent?

Answer. Since X,Y are independent exponential($\lambda=1$), their sum $V = X+Y \sim \Gamma(\lambda=1, n=2)$ with pdf

$$f_V(v) = ve^{-v}, v>0.$$

By part a),

$$f_U(u) = \int_{-\infty}^{\infty} g(u,v)\,dv = I_{\{0<u<1\}} \int_0^{\infty} ve^{-v} dv = I_{\{0<u<1\}},$$

and $g(u,v) = f_U(u) f_V(v), 0<u<1, v>0$: we see that U is uniform in $(0,1)$, and U,V are independent.

c) Show that $\mathbf{P}(rX<Y) = \frac{1}{1+r}, r>0.$

Answer. Since X,Y are independent, for $r>0$,

$$\begin{aligned} \mathbf{P}(rX<Y) &= \int\int_{\{rx<y\}} e^{-x} e^{-y} dy dx = \int_0^{\infty} e^{-x} \left(\int_{rx}^{\infty} e^{-y} dy \right) dx \\ &= \int_0^{\infty} e^{-x} e^{-rx} dx = \frac{1}{1+r} \int_0^{\infty} (1+r) e^{-(1+r)x} dx = \frac{1}{1+r}. \end{aligned}$$

`Other way,` since $X/(X+Y)$ is uniform in $(0,1)$,by part b) with $r>0$,

$$\begin{aligned} \mathbf{P}(rX<Y) &= \mathbf{P}(rX+X<X+Y) = \mathbf{P}((1+r)X < X+Y) \\ &= \mathbf{P}\left(\frac{X}{X+Y} < \frac{1}{1+r} \right) = \frac{1}{1+r}. \end{aligned}$$

5. Let X,Y be independent standard normal random variables.

(a) Find the number a for which $U = X+2Y$ and $V = aX+Y$ are independent.

Answer. Since U,V have zero mean, and X,Y are independent with $\mathbf{E}(X^2) = \mathbf{E}(Y^2) = 1$, $E(X) = \mathbf{E}(Y) = 0$, we have

$$\begin{aligned} \mathrm{Cov}(U,V) &= \mathbf{E}(UV) = \mathbf{E}[(X+2Y)(aX+Y)] \\ &= \mathbf{E}\left(aX^2 + XY + 2aXY + 2Y^2\right) = a+2 = 0 \end{aligned}$$

if $a = -2$. Alternatively, we could compute $\mathrm{Cov}(U, V)$ using bilinearity of covariance and independence of X, Y. Hence U, V are independent if $(U, V) = (X + 2Y, -2X + Y)$ is normal bivariate. It is indeed the case because $(U, V) = (X, Y)\, A$ with

$$A = \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}, \det(A) = 5 \neq 0.$$

(b) Find $\mathbf{E}(X|X + 2Y = z)$, and $\mathbf{E}\left(X^2|X + 2Y = z\right)$ for all $z \in \mathbf{R}$. What is the best mean square estimate of X^2 given $X + 2Y = 1$?

Answer. First note that $(X, X + 2Y)$ is normal bivariate because $(X, X + 2Y) = (X, Y)\, B$ with

$$B = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}, \det(B) = 2 \neq 0.$$

We find its parameters (recall X, Y are independent):

$$\begin{aligned} \mu_1 &= \mathbf{E}(X) = \mathbf{E}(X + 2Y) = \mu_2 = 0, \sigma_1^2 = \mathrm{Var}(X) = 1, \sigma_2^2 = \mathrm{Var}(X + 2Y) = 1 + 4 = 5, \\ \rho &= \frac{\mathrm{Cov}(X, X + 2Y)}{\sigma_1 \sigma_2} = \frac{1}{\sqrt{5}}. \end{aligned}$$

Hence

$$X = \rho \frac{\sigma_1}{\sigma_2}(X + 2Y) + V = \frac{1}{5}(X + 2Y) + V,$$

where $V \sim N\left(0, \sigma_1^2\left(1 - \rho^2\right)\right) = N\left(0, \frac{4}{5}\right)$ is independent of $(X + 2Y)$. So, given $X + 2Y = z$,

$$X = \frac{1}{5}z + V, \ V \sim N\left(0, \frac{4}{5}\right).$$

Thus

$$\begin{aligned} \mathbf{E}(X|X + 2Y = z) &= \frac{1}{5}z, \\ \mathbf{E}\left(X^2|X + 2Y = z\right) &= \mathrm{Var}(X|X + 2Y = z) + (\mathbf{E}(X|X + 2Y = z))^2 \\ &= \frac{4}{5} + \left(\frac{1}{5}z\right)^2 = \frac{4}{5} + \frac{z^2}{25}. \end{aligned}$$

The best mean square estimate of X^2 given $X+2Y = 1$, is $\mathbf{E}\left(X^2|X + 2Y = 1\right) = \frac{4}{5} + \frac{1}{25} = \frac{21}{25}$.

6. a) Let $\alpha \in (0, 2]$, and $T_n = X_1 + X_2 + \ldots + X_n$, where X_i are independent identically distributed continuous random variables with characteristic function $\phi(x) = e^{-|x|^\alpha}, x \in \mathbf{R}$.

Find the number r so that $n^r T_n$ has the same distribution as X_1 (Recall: two random variables have the same distribution iff their characteristic functions coincide).

Show that $\bar{X}_n = n^{-1}T_n = \frac{X_1+...+X_n}{n} \to 0$ in probability as $n \to \infty$ if $\alpha \in (1,2]$ (hint: show that $\bar{X}_n \xrightarrow{D} 0$ as $n \to \infty$ if $\alpha \in (1,2]$).

Answer. We consider characteristic functions. By independence of X_i,

$$\phi_{T_n}(t) = \phi(t)^n = e^{-n|t|^\alpha}, t \in \mathbf{R},$$

and

$$\phi_{n^r T_n}(t) = \phi_{T_n}(n^r t) = \exp\{-n|n^r t|^\alpha\} = \exp\{-nn^{r\alpha}|t|^\alpha\} = \exp\left\{-n^{1+r\alpha}|t|^\alpha\right\}, t \in \mathbf{R}.$$

It coincides with $\phi_{X_1}(t) = \phi(t) = \exp\{-|t|^\alpha\}$ for all t iff $n^{1+r\alpha} = 1$, equivalently, $1 + r\alpha = 0, r = -1/\alpha$.

According to our computation with r,

$$\phi_{\bar{X}_n}(t) = \phi_{n^{-1}T_n}(t) = \exp\left\{-n^{1-\alpha}|t|^\alpha\right\}, t \in R.$$

For $\alpha > 1$, we have $n^{1-\alpha} = \frac{1}{n^{\alpha-1}} \to 0$ as $n \to \infty$, and

$$\phi_{\bar{X}_n}(t) = \exp\left\{-n^{1-\alpha}|t|^\alpha\right\} \to 1 \text{ for all } t.$$

Since the characteristic function of a constant c is $\phi_c(t) = e^{ict}, t \in \mathbf{R}$. We see that $\phi_{\bar{X}_n}(t) \to \phi_0(t) = 1$ for all $t : \bar{X}_n \xrightarrow{D} 0$.

b) Let S_n be the number of heads in n tosses of a coin whose heads probability is 0.1. Approximate the probability

$$\mathbf{P}\left(\frac{S_n}{n} \geq 0.14\right)$$

using the distribution function $\Phi(x)$ of a standard normal random variable. Is this probability larger with $n = 100$ or $n = 10000$?

Answer. We know S_n is binomial$(n, p = 0.1)$. According to CLT for binomial(n, p) for large n,

$$\begin{aligned}
\mathbf{P}\left(\frac{S_n}{n} \geq 0.14\right) &= \mathbf{P}\left(\frac{\frac{S_n}{n} - 0.1}{\sqrt{0.1(1-0.1)}/\sqrt{n}} \geq \sqrt{n}\frac{0.14-0.1}{\sqrt{0.1(1-0.1)}}\right) \\
&\approx \mathbf{P}\left(Z \geq \sqrt{n}\frac{0.04}{0.3} = \frac{4}{30}\sqrt{n}\right) = 1 - \Phi\left(\frac{4\sqrt{n}}{30}\right).
\end{aligned}$$

The probability is larger with $n = 100$ than with $n = 10000$: the interval for the standard normal Z is larger with $n = 100$.