1. *Suppose that 20 percent of smokers get lung cancer and 1 percent of non-smokers get lung cancer. In a population there are equal number of males and females, however 30 percent of males are smokers and 20 percent of females are smokers. A person was chosen at random from the population. What is the probability that he/she has lung cancer? If the chosen person has lung cancer, what is the conditional probability that it is a non-smoking male?*

   **Solution:** Let $C$ be the event that a randomly chosen person has lung cancer,
   $S$ be the event that a randomly chosen person is a smoker and
   $M$ be the event that a randomly chosen person is a male.
   Then, we are given: $P(C/S) = 20\% = 0.2$, $P(C/S^C) = 1\% = 0.01$ and $P(S/M) = 30\% = 0.3$ and $P(S/M^C) = 20\% = 0.2$.
   Now, $\frac{P(C \cap S)}{P(S)} = 0.2$ and $\frac{P(C \cap S^C)}{P(S^C)} = 0.01$
   Also, $\frac{P(S \cap M)}{P(M)} = 0.3$ and $\frac{P(S \cap M^C)}{P(M^C)} = 0.2$
   Since it is given that there are equal number of males and females, so $P(M) = P(M^C) = 0.5$, $\therefore$ $P(S \cap M) = 0.3 \times 0.5 = 0.15$ and $P(S \cap M^C) = 0.2 \times 0.5 = 0.1$
   $P(S) = P(S \cap M) + P(S \cap M^C) = 0.25$ and $P(S^C) = 0.75$ Thus, $P(S \cap C) = 0.2 \times 0.25 = 0.05$ and $P(C \cap S^C) = 0.01 \times 0.75 = 0.0075$ and hence,

   $$P(C) = 0.005 + 0.0075 = 0.0575$$

   For the next part, we need to compute $P(S^C \cap M/C)$.
   Consider

   $$\begin{aligned}
   & P(C \cap S^C \cap M) \\
   = \;& P(M \cap S^C)\ .\ P(C/M \cap S^C) \\
   = \;& P(M \cap S^C)\ .\ P(C/M^C \cap S^C) \\
   = \;& P(M \cap S^C)\ .\ P(C/S^C), \text{ (as a non-smoker gets cancer with probability 0.01 irrespective of whether} \\
   & \text{the person is male or female)} \\
   = \;& 0.35 \times 0.01 = 0.0035
   \end{aligned}$$

   Thus, $P(S^C \cap M/C) = \frac{P(S^C \cap C \cap M)}{P(C)} = \frac{35}{575} = 0.06$, which is the required probability. ☐

2. *Let $k$ be a natural number. Consider a coin, where the chance of Head is $p$ with $0 < p \leqslant 1$. Let $G$ be the number of independent tosses of the coin to obtain $k$ many Heads in succession (i.e., consecutively) for the first time. Show that $P(G = k+1) = P(G = k+2) = \cdots = P(G = 2k)$.*

   **Solution:**

   $$\begin{aligned}
   P(G = k+1) &= P(\text{Tail at first toss and then all Heads}) \\
   &= p^k\ (1-p),
   \end{aligned}$$

   $$\begin{aligned}
   P(G = k+2) &= P(\text{Head or Tail at first toss, Tail at second toss and then all Heads}) \\
   &= p\ (1-p)\ p^k + (1-p)\ (1-p)\ p^k \\
   &= p^k\ (1-p),
   \end{aligned}$$

$$\vdots$$

$$\begin{aligned}
P(G = 2k) &= P(\text{Head or Tail at } (k-1) \text{ places, Tail at } k\text{-th place and then } k \text{ successive Heads}) \\
&= \sum_{s \in \{H,T\}^{k-1}} P(s, T, HHH \dots H) \\
&= \left[ \sum_{s \in \{H,T\}^{k-1}} P(s) \right] (1-p)\ p^k \\
&= 1\ (1-p)\ p^k = p^k\ (1-p).
\end{aligned}$$

Thus, we get the desired conclusion. □

3. *Suppose $X, Y$ are independent random variables. Assume that $X$ has Binomial distribution with parameters $(m, p)$ and $Y$ has Binomial distribution with parameters $(n, p)$, where $m, n$ are natural numbers and $0 \leqslant p \leqslant 1$. Show that $Z = X+Y$ has Binomial distribution with parameters $(m+n, p)$. Show that $X - Y$ does not have Binomial distribution, even if $m \geqslant n$.*

**Solution:** The m.g.f of $X$ is given by:

$$\begin{aligned}
E[e^{Xt}] &= \sum_{j=0}^{m} e^{jt} \binom{n}{j} p^j\ (1-p)^{m-j}, \quad \forall t \in \mathbb{R} \\
&= \sum_{j=0}^{m} \binom{n}{j} (p\ e^t)^j\ (1-p)^{m-j} \\
&= \Big( (1-p) + p\ e^t \Big)^m.
\end{aligned}$$

Similarly, the m.g.f of $Y$ is $E[e^{Yt}] = \Big( (1-p) + p\ e^t \Big)^n$.

The m.g.f of $Z = X + Y$ is given by

$$\begin{aligned}
E[e^{(X+Y)t}] &= E[e^{Xt}\ .\ e^{Yt}] \\
&= E[e^{Xt}]\ E[e^{Yt}] \\
&= \Big( (1-p) + p\ e^t \Big)^{m+n}.
\end{aligned}$$

Hence, $Z = X + Y$ has Binomial distribution with parameters $(m + n, p)$.

When $m < n$, it is possible for $X - Y$ to be negative, which is not the case for any binomially distributed random variables.

When $m \geqslant n$,

$$\begin{aligned}
E[e^{(X-Y)t}] &= E[e^{Xt}\ .\ e^{-Yt}] \\
&= E[e^{Xt}]\ E[e^{Y(-t)}] \\
&= \Big( (1-p) + p\ e^t \Big)^m \Big( (1-p) + p\ e^{-t} \Big)^n,
\end{aligned}$$

which shows that $X - Y$ just cannot be expressed as being a count of successes in a sequence of some amount of independent, identically distributed Bernoulli trials. □

4. *Suppose $U$ is a random variable having uniform distribution in the interval $[-2,2]$. Compute probability distribution function and densities of $R = \frac{U}{2} + 1$ and $S = U^2 + 1$.*

**Solution:** Since $U \sim Unif[-2,2]$, so we have

$$f_U(u) = \begin{cases} \frac{1}{4} & \forall u \in [-2,2] \\ 0 & \text{otherwise} \end{cases},$$

$R = \frac{U}{2} + 1 \in [0,2]$ and $S = U^2 + 1 \in [0,5]$.
Also, we have $U = 2(r-1)$, so

$$\begin{aligned} f_R(r) &= f(2(r-1)) \times 2, \quad \text{if} \quad 2(r-1) \in [-2,2], \ i.e., \ r \in [0,2] \\ &= \frac{1}{4} \times 2 = \frac{1}{2}, \end{aligned}$$

$\therefore$ the density is given by

$$f_R(r) = \begin{cases} \frac{1}{2} & \forall r \in [0,2] \\ 0 & \text{otherwise} \end{cases},$$

and probability density function of $R$ is given by

$$F_R(r) = \int_0^r f_R(t)\ dt = \int_0^r \frac{1}{2}\ dt = \frac{r}{2}.$$

Similarly, the density of $S$ is given by

$$f_S(s) = \begin{cases} \frac{1}{4\sqrt{s-1}} & \forall s \in [0,5] \\ 0 & \text{otherwise} \end{cases},$$

and probability density function of $S$ is given by

$$F_S(s) = \int_0^s f_S(t)\ dt = \int_0^s \frac{1}{4\sqrt{t-1}}\ dt = \frac{\sqrt{s-1}}{2}.$$

□

5. *Suppose $X$ has Binomial distribution with parameters $(n,p)$, where $n$ is a natural number and $0 < p < 1$. Suppose $Y$ is a random variable defined as a function of $X$ by*

$$Y = \begin{cases} 0 & \textit{if} \quad X = 0 \\ 1 & \textit{otherwise} \end{cases}$$

*Compute the conditional distribution of $X$ given $Y = 1$ and the conditional expectation of $X$ given $Y = 1$.*

**Solution:** The conditional distribution:

$$\begin{aligned} P(X = k/Y = 1) &= \frac{P(X = k \cap Y = 1)}{P(Y = 1)} \\ &= \frac{P(X = k \cap X \neq 0)}{P(Y = 1)}. \end{aligned}$$

Now, if $k = 0$, then $P(X = k/Y = 1) = 0$.
If $k \neq 0$, then

$$\begin{aligned}
\frac{P(X=k)}{P(Y=1)} &= \frac{P(X=k)}{P(X \neq 0)} \\
&= \frac{P(X=k)}{1 - P(X=0)} \\
&= \frac{P(X=k)}{1 - \binom{n}{0}\ p^0\ (1-p)^{n-0}} \\
&= \frac{P(X=k)}{1-(1-p)^n}.
\end{aligned}$$

Next, the conditional expectation:

$$\begin{aligned}
E[X/Y=1] &= \sum_{j=0}^{n} j\ P(X=j/Y=1) \\
&= \sum_{j=1}^{n} j\ P(X=j/Y=1) \\
&= \frac{1}{1-(1-p)^n} \sum_{j=1}^{n} j \binom{n}{j}\ p^j\ (1-p)^{n-j} \\
&= \frac{np}{1-(1-p)^n} \sum_{j=1}^{n} \frac{(n-1)!}{(n-j)!\ (j-1)!}\ p^{j-1}\ (1-p)^{n-j} \\
&= \frac{np}{1-(1-p)^n} \sum_{j=1}^{n} \binom{n-1}{j-1}\ p^{j-1}\ (1-p)^{n-j} \\
&= \frac{np}{1-(1-p)^n}\ (p + (1-p))^{n-1} \\
&= \frac{np}{1-(1-p)^n}.
\end{aligned}$$

□

6. *Suppose $\lambda > 0$ and $E$ is an exponential random variable with parameter $\lambda$. Compute a probability density function for $K = 1 + \sqrt{E}$.*

**Solution:** For any $k \geqslant 0$, the cumulative distribution function is given by

$$\begin{aligned}
P(K \leqslant k) &= P(1+\sqrt{E} \leqslant k) \\
&= P(\sqrt{E} \leqslant k-1) \\
&= P(E \leqslant (k-1)^2), \quad \text{as exponential random variable is supported on } [0,\infty) \\
&= 1 - e^{-\lambda\ (k-1)^2}.
\end{aligned}$$

Thus, the probability density function of $K$ is

$$f_K(k) = 2\ \lambda\ (k-1)\ e^{-\lambda\ (k-1)^2}.$$

□