Possible Solutions

Probability Theory and Statistics (BSc D-ITET)

1. a) Using the same notation as in the exercise, we know that

$$P(\nabla^c \mid G) = 0.2, \qquad P(M \cap \nabla) = 0.5, \qquad \text{and} \qquad P(G) = 0.3.$$

i) Since $M = G^c$, we know that $P(M) = 1 - P(G) = 0.7$. Using the definition of conditional probabilities

$$P(\nabla \mid M) = \frac{P(M \cap \nabla)}{P(M)} = \frac{0.5}{0.7} \ (\approx 0.714).$$

ii) From the definition of conditional probabilities

$$P(\nabla \cap G) = P(\nabla \mid G)P(G).$$

Observe that $P(\nabla \mid G) = 1 - P(\nabla^c \mid G) = 0.8$, hence

$$P(\nabla \cap G) = 0.24.$$

Remark: The same computation, doing all the steps reads:

$$P(\nabla \mid G) = \frac{P(\nabla \cap G)}{P(G)} = \frac{P(G) - P(\nabla^c \cap G)}{P(G)} = \frac{P(G)(1 - P(\nabla^c \mid G))}{P(G)} = 1 - P(\nabla^c \mid G).$$

However it is not necessary to prove this formula.

b) We can solve it using Bayes theorem or the definition of conditional probabilities. In any case we need to compute $P(\nabla)$. Using the definition of conditional expectation, since

$$P(\nabla) = P(M \cap \nabla) + P(G \cap \nabla) = 0.5 + 0.24 = 0.74.$$

we get

$$P(M \mid \nabla) = \frac{P(\nabla \cap M)}{P(\nabla)} = \frac{0.5}{0.74} \ (\approx 0.676).$$

Alternative solution: We know that $P(\nabla \mid M) = \frac{0.5}{0.7}$ and $P(\nabla \mid G) = 0.8$. Therefore, from Bayes theorem, we get

$$P(M \mid \nabla) = \frac{P(\nabla \mid M)P(M)}{P(M)P(\nabla \mid M) + P(G)P(\nabla \mid G)} = \frac{\frac{0.5}{0.7} \cdot 0.7}{0.7 \cdot \frac{0.5}{0.7} + 0.3 \cdot 0.8} = \frac{0.5}{0.74}.$$

Please turn the page!

c) We denote by X_s the random variable counting the number of people in line for the sorbets at 3 PM, and by X_c the one counting the number of people in line for the creamy ice–creams at 3 PM. We set $X = X_s + X_c$.

i) We calculate then (with $\lambda = \lambda_c + \lambda_s$)

$$\begin{aligned} P(X_s = k \mid X = 20) &= \frac{P(X_s = k, X = 20)}{P(X = 20)} = \frac{P(X_s = k, X_c = 20 - k)}{P(X = 20)} \\ &\overset{indep.}{=} \frac{P(X_s = k)P(X_c = 20 - k)}{P(X = 20)} = \frac{e^{-\lambda_s}\lambda_s^k}{k!} \frac{e^{-\lambda_c}\lambda_c^{(20-k)}}{(20-k)!} \frac{20!}{e^{-\lambda}\lambda^{20}} \\ &= \binom{20}{k} \left(\frac{\lambda_s}{\lambda}\right)^k \left(\frac{\lambda_c}{\lambda}\right)^{20-k} = \binom{20}{k} (0.25)^k (0.75)^{20-k} \end{aligned}$$

We recognize that this is the distribution of a Binomial distribution with parameters $n = 20$ and $p = 0.25$.

ii) We know from the lecture that the expectation of a Binomial random variable with parameters n and p is given by np. In this case

$$E[X_s \mid X = 20] = 20 \cdot 0.25 = 5.$$

2. **a)** The marginal density of Y is then given, for $y > 0$, by

$$f_Y(y) = \int_{-\infty}^{+\infty} f_{X,Y}(x, y)dx = \int_y^{+\infty} \frac{1}{2} x e^{-x} dx \overset{i.b.p}{=} \frac{1}{2}(y + 1)e^{-y}.$$

Hence we get $f_Y(y) = \frac{1}{2}(y + 1)e^{-y}\, 1(y > 0)$.

b) Let us calculate

$$E[Y] = \int_{-\infty}^{+\infty} y f_Y(y)dy = \int_0^{+\infty} \frac{1}{2} y(y + 1)e^{-y}dy \overset{hint}{=} \frac{3}{2}.$$

c) Using that

$$E[XY] = \int_0^{\infty} \int_y^{\infty} \frac{1}{2} x^2 y e^{-x} dx dy = \int_0^{\infty} \left(\int_0^x y dy\right) \frac{1}{2} x^2 e^{-x} dx = 6,$$

part **b)**, and the fact that $E[X] = 3$ (which can be obtained by the hint in part **b)**), we conclude that

$$\mathrm{Cov}(X, Y) = E[XY] - E[X]E[Y] = \frac{3}{2}.$$

d) The random variables X, Y are indeed dependent. Two acceptable reasons for this are:

i) The covariance of X and Y is non-zero.

ii) The joint density function is not the product of the marginal densities.

See next page!

3. **a)** $\{T = n\}$ means exactly that $D_i \neq 1$ for $i = 1, \dots, n-1$, and $D_n = 1$. In set theoretic language,

$$\{T = n\} = \left(\bigcap_{i=1}^{n-1} \{D_i \neq 1\} \right) \cap \{D_n = 1\}.$$

Additionally, since $\{D_i \neq 1\} = \{D_i = 1\}^c$, we obtain

$$P(D_i \neq 1) = 1 - P(D_i = 1) = 1 - P(D = 1) = 1 - p.$$

Using the fact that the D_i are independent, we get

$$P(T = n) = P\left[\left(\bigcap_{i=1}^{n-1} \{D_i \neq 1\} \right) \cap \{D_n = 1\} \right] = \left(\prod_{i=1}^{n-1} P(D_i \neq 1) \right) \times P(D_n = 1)$$
$$= (1-p)^{(n-1)} p.$$

T has then a geometric distribution with parameter p.

b) Using part **a)**, we get

$$E[T] = \sum_{n=1}^{\infty} nP(T = n) \overset{\mathbf{a)}}{=} \sum_{n=1}^{\infty} n(1-p)^{(n-1)} p = p \sum_{n=1}^{\infty} n(1-p)^{(n-1)}.$$

We have also the identity seen in class for $z \in [0, 1)$,

$$\sum_{n=1}^{\infty} nz^{(n-1)} = \frac{d}{dz} \left(\frac{1}{1-z} \right) = \frac{1}{(1-z)^2}.$$

Using the substitution $z = 1 - p \in (0, 1)$, we get

$$E[T] = p \sum_{n=1}^{\infty} n(1-p)^{(n-1)} = p \cdot \frac{1}{p^2} = \frac{1}{p}.$$

Alternative solution: Using part **a)**, we get

$$E[T] = \sum_{n=1}^{\infty} nP(T = n) \overset{\mathbf{a)}}{=} \sum_{n=1}^{\infty} n(1-p)^{(n-1)} p = \lim_{N\to\infty} p \sum_{n=1}^{N} n(1-p)^{(n-1)}.$$

Setting $I_N = \sum_{n=1}^{N} n(1-p)^{(n-1)}$, we calculate

$$(1-p)I_N = \sum_{n=1}^{N} n(1-p)^n \;\Rightarrow\; pI_N = \left(\sum_{n=1}^{N} (1-p)^{(n-1)} \right) - N(1-p)^N.$$

Furthermore, using $\sum_{n=1}^{N} (1-p)^{(n-1)} = \frac{1-(1-p)^N}{1-(1-p)} = \frac{1-(1-p)^N}{p}$ for $p \in (0, 1)$, we obtain

$$I_N = \frac{1-(1-p)^N}{p^2} - \frac{N(1-p)^N}{p} \;\Rightarrow\; E[T] = \lim_{N\to\infty} pI_N = \lim_{N\to\infty} \left(\frac{1-(1-p)^N}{p} - N(1-p)^N \right).$$

Since $(1-p) \in (0, 1)$, we have $\lim_{N\to\infty}(1-p)^N = \lim_{N\to\infty} N(1-p)^N = 0$. Therefore, we get $E[T] = \frac{1}{p}$.

Please turn the page!

c) Using the standard approximation $E[T] \approx \frac{1}{k}\sum_{i=1}^{k} T_i$, we can get an estimator for p using the identity proven in part **b)**, i.e. $p = \frac{1}{E[T]}$,

$$\widehat{p} = \left(\frac{1}{k}\sum_{i=1}^{k} T_i\right)^{-1}.$$

d) Since T_i has also a geometric distribution with parameter p, we take $p_{T_i}(t_i) = (1-p)^{(t_i-1)}p$ for $i = 1, \dots, k$ (discrete distribution). Therefore, using the independence of T_i with $i = 1, \dots, k$, the likelihood function $L(p, t_1, \dots t_k)$ is given by

$$L(p, t_1, \dots t_k) = \prod_{i=1}^{k} p_{T_i}(t_i) = \prod_{i=1}^{k} (1-p)^{(t_i-1)}p.$$

Hence, the log-likelihood is given by

$$\ell(p, t_1, \dots t_k) = \log L(p, t_1, \dots t_k) = \sum_{i=1}^{k} [(t_i - 1)\log(1-p) + \log p] = k \log p + \sum_{i=1}^{k} (t_i - 1)\log(1-p).$$

To maximize the log likelihood function, we set the derivative with respect to p equal to 0, i.e.

$$\frac{\partial}{\partial p}\ell(p, t_1, \dots t_k) = 0 \Leftrightarrow \frac{k}{p} + \sum_{i=1}^{k}(t_i - 1)\frac{-1}{1-p} = 0 \Leftrightarrow k(1-p) = p\sum_{i=1}^{k}(t_i - 1)$$

Since $\sum_{i=1}^{k}(t_i - 1) = -k + \sum_{i=1}^{k} t_i$, we solve for p (and replacing each realization t_i with its associated random variable T_i) to get the maximum-likelihood estimator

$$\widehat{p} = \left(\frac{1}{k}\sum_{i=1}^{k} T_i\right)^{-1}.$$

Remark: The maximum-likelihood estimator is the same as the "natural" estimator of part **c)**.