King Saud University
College of Sciences
Mathematics Department

Academic Year (G) 2016–2017
Academic Year (H) 1437–1438
Bachelor AFM: M. Eddahbi

Solutions of Homework 1 : Selected problems P exam

Problem 1:
An auto insurance company has 10000 policyholders. Each policyholder is classified as:
(i) young or old;
(ii) male or female;
(iii) married or single.

Of these policyholders, 3000 are young, 4600 are male and 7000 are married. The policyholders can also be classified as 1320 young males, 3010 married males and 1400 young married persons. Finally, 600 of the policyholders are young married males.

Question: How many of the company's policyholders are young, female and single?

Solution:

	Classification			
	Young	Married	Single	Total
Number	3000	7000	3000	10000
Males	1320	3010	1590	4600
Females	1680	3990	1410	5400
	Young	Young and married	Young and single	Total
Number	3000	1400	1600	3000
Males	1320	600	$1320 - 600 = 720$	1320
Females	1680	800	$1680 - 800 = 880$	1680

The number of policyholders who are young, female and single is 880.

Problem 2.
A survey of a group's viewing habits over the last year revealed the following information
(i) 28% watched gymnastics
(ii) 29% watched baseball
(iii) 19% watched soccer
(iv) 14% watched gymnastics and baseball
(v) 12% watched baseball and soccer
(vi) 10% watched gymnastics and soccer
(vii) 8% watched all three sports.

Question: Calculate the percentage of the group that watched none of the three sports during the last year.

Choose one of the following answers: (A) 24, (B) 36, (C) 41, (D) 52, (E) 60.

Solution: (D) 52%. Because the percentage of the group that watched none of the three sports during the last year is equal to 100 minus the percentage of the group that watched at least one of the three sports during the last year which is equal to $100 - 28 - 29 - 19 + 14 + 12 + 10 - 8 = 52$.

More precisely denote by p the probability that the group watched none of the three sports. Let us denote by WG: watched gymnastics, WB : watched baseball and WS watched soccer. Then

$$p = 1 - P\left(WG \ \ or \ WB \ or \ WS\right).$$

First we have

$$\begin{aligned}
P\left(WG \ \ or \ WB \ or \ WS\right) &= P\left(WG\right)+P\left(WB\right)+P\left(WS\right) \\
&\quad -P\left(WG \ \ and \ WB\right)-P\left(WB \ and \ WS\right)-P\left(WG \ and \ WS\right) \\
&\quad +P\left(WG \ \ and \ WB \ and \ WS\right) \\
&= 0.28+0.29+0.19-0.14-0.12-0.10+0.08=0.48
\end{aligned}$$

Therefore $p=1-P\left(WG \ \ or \ WB \ or \ WS\right)=0.52=52\%$.

Problem 3.

An insurance company estimates that 40% of policyholders who have only an auto policy will renew next year and 60% of policyholders who have only a homeowners policy will renew next year. The company estimates that 80% of policyholders who have both an auto and a homeowners policy will renew at least one of those policies next year. Company records show that 65% of policyholders have an auto policy, 50% of policyholders have a homeowners policy, and 15% of policyholders have both an auto and a homeowners policy.
Question: Using the company's estimates, calculate the percentage of policyholders that will renew at least one policy next year.
Solution: Let us denote by A event that a policyholder has an auto policy, H stands for the event that a policyholder has a homeowners policy and $A\cap H$ the event that a policyholder has both an auto and a homeowners policy.

Denote by R the event that a policyholder renews at least one policy.

The proportion of policyholders that will renew at least one policy is given by $P(R)$. Remark first that

$$(A\cap H)\cup(A\cap H^c)\cup(A^c\cap H)=\Omega=\text{the whole space}$$

and the sets $(A\cap H)$, $(A\cap H^c)$ and $(A^c\cap H)$ are disjoint, therefore

$$P(R)=P\left(R\cap\left((A\cap H)\cup(A\cap H^c)\cup(A^c\cap H)\right)\right).$$

Now, if we set $A\cap H=A_1$, $A\cap H^c=A_2$ and $A^c\cap H=A_3$

$$\begin{aligned}
P(R) &= P\left(R\cap(A_1\cup A_2\cup A_3)\right)=P\left(R\cap A_1\cup R\cap A_2\cup R\cap A_3\right) \\
&= P\left(R\cap A_1\right)+P\left(R\cap A_2\right)+P\left(R\cap A_3\right) \\
&= P(A_1)P(R\mid A_1)+P(A_2)P(R\mid A_2)+P(A_3)P(R\mid A_3).
\end{aligned}$$

From the given information we can calculate $P(A_1)=P(A\cap H)$, $P(A_2)=P(A\cap H^c)$ and $P(A_3)=P(A^c\cap H)$.

$$\begin{aligned}
P(A\cap H) &= 0.15, \\
P(A\cap H^c) &= P(A)-P(A\cap H)=0.65-0.15=0.50 \\
P(A^c\cap H) &= P\left(H\right)-P(A\cap H)=0.50-0.15=0.35
\end{aligned}$$

Consequently

$$P(R)=50\%\times 40\%+35\%\times 60\%+15\%\times 80\%=53\%.$$

Problem 4.

An urn contains N balls numbered from 1 to N. We pick a randomly a ball (all the balls are equally likely to be extracted) and define the r.v. X by the number of the extracted ball.

1. Find the distribution of X.
   a. The sample space S_X or S of X is the set of all possible values of X which is given by $S_X = \{1, 2, 3, \ldots, N\}$.
   b. For all $k \in S_X$ we need the p.m.f. $f(k)$. In our case we have $f(k) = \frac{1}{N}$ for all $k \in S_X$. Then the distribution of X is given by $(k, \frac{1}{N})$ $k \in S_X$. This is a **uniform distribution** in S_X.

2. Let X be a r.v. with values in $\mathbb{N}$ such that: $\forall\, n \in \mathbb{N}^*$, $P(X = n) = \dfrac{a}{n} P(X = n-1)$ where $a > 0$.

   (a) Find $P(X = 0)$
   Solution: We have

$$\begin{aligned} P(X = n) &= \frac{a}{n} P(X = n-1) = \frac{a}{n}\frac{a}{n-1}\cdots\frac{a}{2}\frac{a}{1} P(X = 0) \\ &= \frac{a^n}{n!} P(X = 0) \quad \forall\, n \in \mathbb{N}^*. \end{aligned}$$

   We know that

$$\begin{aligned} \sum_{n=0}^{\infty} P(X = n) &= 1 \iff P(X = 0) + \sum_{n=1}^{\infty} \frac{a^n}{n!} P(X = 0) = 1 \\ &\iff P(X = 0)\left(1 + \sum_{n=1}^{\infty} \frac{a^n}{n!}\right) = P(X = 0)\left(\sum_{n=0}^{\infty} \frac{a^n}{n!}\right) = 1 \\ &\iff P(X = 0) = e^{-a} \end{aligned}$$

   (b) Find the distribution of X.
   Solution: We have $S_X = \mathbb{N}$ and for all $n \in \mathbb{N}$, the p.m.f. $f(n) = P(X = n) = e^{-a}\frac{a^n}{n!}$. So X follows a Poisson distribution with paramter $a > 0$.

3. Let p_n, $n = 0, 1, 2, \ldots$, be the probability that an automobile policyholder will file for n claims in a five–year period. The actuary involved makes the assumption that $p_{n+1} = \frac{1}{4} p_n$.

   (a) What is the probability that the holder will file two or more claims during this period?
   Solution: Let X be the number of filed claims by automobile policyholder in a five–year period. We have $p_n = P(X = n)$. We want to caclulate $P(X \geq 2)$. We can write

$$\begin{aligned} P(X \geq 2) &= 1 - P(X \leq 1) = 1 - p_0 - p_1 \\ &= 1 - p_0 - \frac{1}{4} p_0 = \end{aligned}$$

   Now we need to compute p_0. We have

$$\begin{aligned} \sum_{n=0}^{\infty} p_n &= 1 \iff p_0 + \frac{1}{4} p_0 + \frac{1}{4^2} p_0 + \cdots + \frac{1}{4^n} p_0 + \cdots = 1 \\ &\iff p_0 \left(\sum_{n=0}^{\infty} \frac{1}{4^n}\right) = 1 \iff p_0 \frac{1}{1 - \frac{1}{4}} = 1 \iff p_0 \frac{4}{3} = 1 \end{aligned}$$

   Hence $p_0 = \frac{3}{4}$. Finaly

$$P(X \geq 2) = 1 - \frac{5}{4} \times \frac{3}{4} = 1 - \frac{15}{16} = \frac{1}{16}.$$

Problem 5.

1. An insurance policy pays 100 per day for up to three days of hospitalization and 50 per day for each day of hospitalization thereafter. The number of days of hospitalization, X, is a discrete random variable with probability function

$$P(X=k)=\begin{cases} \dfrac{6-k}{15} & \text{for } k=1,2,3,4,5 \\ 0 & \text{otherwise.} \end{cases}$$

   (a) Determine the expected payment for hospitalization under this policy.
   Solution: Let Z be the r.v. corresponding the payment for hospitalization under this policy. We have

$$\begin{aligned} E[Z] &= 100\times P(X=1)+100\times 2\times P(X=2)+100\times 3\times P(X=3) \\ &\quad +50\times 4\times P(X=4)+50\times 5\times P(X=5). \end{aligned}$$

$$E[Z]=100\times\frac{1}{3}+100\times 2\times\frac{4}{15}+100\times 3\times\frac{1}{5}+50\times 4\times\frac{2}{15}+50\times 5\times\frac{1}{15}=190.$$

2. The number of injury claims per month is modeled by a random variable X with

$$P(X=n)=\frac{1}{(n+1)(n+2)}$$

for nonnegative integers, n.

   (a) Calculate the probability of at least one claim during a particular month, given that there have been at most four claims during that month.
   Solution:

$$\begin{aligned} P(X\geq 1\mid X\leq 4) &= \frac{P(X\geq 1,\ X\leq 4)}{P(X\leq 4)}=\frac{P(1\leq X\leq 4)}{P(X\leq 4)} \\ &= \frac{P(X\leq 4)=P(X=0)}{P(X\leq 4)}=1-\frac{P(X=0)}{P(X\leq 4)}. \end{aligned}$$

   We have $P(X=1)=\frac{1}{2}$ and

$$\begin{aligned} P(X\leq 4) &= \sum_{n=0}^{4}P(X=n)=\sum_{n=0}^{4}\frac{1}{(n+1)(n+2)} \\ &= \sum_{n=0}^{4}\left(\frac{1}{n+1}-\frac{1}{n+2}\right)=\frac{5}{6}. \end{aligned}$$

   Hence

$$P(X\geq 1\mid X\leq 4)=1-\frac{\frac{1}{2}}{\frac{5}{6}}=1-\frac{1}{2}\frac{6}{5}=\frac{2}{5}.$$

   Other way: You can also write

$$\begin{aligned} P(X\geq 1\mid X\leq 4) &= \frac{P(X\geq 1,\ X\leq 4)}{P(X\leq 4)}=\frac{P(1\leq X\leq 4)}{P(X\leq 4)} \\ &= \frac{P(1\leq X\leq 4)}{P(X=0)+P(1\leq X\leq 4)}. \end{aligned}$$

   But we have

$$P(1\leq X\leq 4)=\sum_{n=1}^{4}\left(\frac{1}{n+1}-\frac{1}{n+2}\right)=\frac{1}{2}-\frac{1}{6}=\frac{2}{6}=\frac{1}{3}.$$

Hence

$$P(X \geq 1 \mid X \leq 4) = \frac{\frac{1}{3}}{\frac{1}{2}+\frac{1}{3}} = \frac{2}{5}.$$

3. An actuary has discovered that policyholders are three times as likely to file two claims as to file four claims. The number of claims filed has a Poisson distribution.

   (a) Calculate the variance of the number of claims filed.
   Solution: Let X be the random number of claims filed by policyholders. We know that the variance of a Poisson distribution is λ. But we have $3P(X=4) = P(X=2)$ that is

$$3e^{-\lambda}\frac{\lambda^4}{4!} = e^{-\lambda}\frac{\lambda^2}{2!} \iff \lambda^2 = \frac{1}{3}\frac{4!}{2!} = 4 \iff \lambda = 2$$

   hence the variance of X is equal to 2.

4. A company establishes a fund of 120 from which it wants to pay an amount, C, to any of its 20 employees who achieve a high performance level during the coming year. Each employee has a 2% chance of achieving a high performance level during the coming year. The events of different employees achieving a high performance level during the coming year are mutually independent.

   (a) Calculate the maximum value of C for which the probability is less than 1% that the fund will be inadequate to cover all payments for high performance.
   Solution: Let X denote the number of employees that achieve the high performance level. It is clear the X follows a binomial distribution with parameters $n = 20$ and $p = 2\%$. We are asked to find m such that

$$P(X > m) \leq 1\% \quad \text{and} \quad mC = 120.$$

   Now, we can write

$$\begin{aligned} P(X > m) &\leq 1\% \iff P(X \leq m) > 99\% \iff \sum_{k=0}^{m} P(X = k) > 99\% \\ &\iff \sum_{k=0}^{m} \frac{20!}{k!(20-k)!}\left(\frac{2}{100}\right)^k\left(\frac{98}{100}\right)^{20-k} > 99\%. \end{aligned}$$

   Now take different values of m,
   for $m = 1$, we have $\sum_{k=0}^{1} \frac{20!}{k!(20-k)!}\left(\frac{2}{100}\right)^k\left(\frac{98}{100}\right)^{20-k} = 0.0173$.
   for $m = 2$, we have $\sum_{k=0}^{2} \frac{20!}{k!(20-k)!}\left(\frac{2}{100}\right)^k\left(\frac{98}{100}\right)^{20-k} = 0.992\,93 > 99\%$.
   Therefore $2C = 120$ then $C = 60$.

5. The loss due to a fire in a commercial building is modeled by a random variable X with density function

$$f(x) = \begin{cases} 0.005(20-x) & \text{for } 0 < x < 20 \\ 0 & \text{otherwise} \end{cases}$$

   (a) Given that a fire loss exceeds 8, calculate the probability that it exceeds 16.
   Solution: We want to compute $P(X > 16 \mid X > 8)$. So we can write by definition

$$\begin{aligned} P(X > 16 \mid X > 8) &= \frac{P(X > 16;\ \ X > 8)}{P(X > 8)} \\ &= \frac{P(X > 16)}{P(X > 8)} \text{ (since } \{X > 16\} \subset \{X > 8\}\text{)} \\ &= \frac{\int_{16}^{20} 0.005(20-x)dx}{\int_{8}^{20} 0.005(20-x)dx}, \end{aligned}$$

But we have $\int_{16}^{20} 0.005(20-x)dx = 0.04$ and $\int_{8}^{20} 0.005(20-x)dx = 0.36$, then

$$P\left(X > 16 \mid X > 8\right) = \frac{4}{36} = \frac{1}{9}.$$

Problem 6.

1. The lifetime of a machine part has a continuous distribution on the interval $]0, 40[$ with probability density function $f(x)$, where $f(x)$ is proportional to $(10+x)^{-2}$ on the interval.

   (a) Calculate the probability that the lifetime of the machine part is less than 6.
   Solution: We want to compute $P\left(T < 6\right)$. So we can write by definition

   $$P\left(T < 6\right) = \int_0^6 \frac{cdx}{(10+x)^2} = \left[-\frac{c}{x+10}\right]_0^6 = \frac{3}{80}c.$$

   Now we have to find c. So since f is a p.d.f. we have

   $$\int_0^{40} \frac{cdx}{(10+x)^2} = 1 \Longleftrightarrow \frac{2}{25}c = 1 \Longleftrightarrow c = \frac{25}{2}.$$

   Therefore

   $$P\left(T < 6\right) = \frac{3}{80} \times \frac{25}{2} = \frac{15}{32}.$$

2. An insurance company insures a large number of homes. The insured value, X, of a randomly selected home is assumed to follow a distribution with density function

   $$f(x) = \begin{cases} 3x^{-4} \text{ for } x > 1 \\ 0 \qquad\qquad\qquad \text{otherwise} \end{cases}$$

   (a) Given that a randomly selected home is insured for at least 1.5, calculate the probability that it is insured for less than 2.
   Solution: We want to compute $P\left(X < 2 \mid X > 1.5\right)$. So we can write by definition

   $$\begin{aligned} P\left(X < 2 \mid X > 1.5\right) &= \frac{P\left(X > 1.5;\ X < 2\right)}{P\left(X > 1.5\right)} = \frac{P\left(1.5 < X < 2\right)}{\left(X > 1.5\right)} \\ &= \frac{\int_{1.5}^{2} 3x^{-4}dx}{\int_{1.5}^{\infty} 3x^{-4}dx} = \frac{\left[-\frac{1}{x^3}\right]_{1.5}^{2}}{\left[-\frac{1}{x^3}\right]_{1.5}^{\infty}} = \frac{\frac{37}{216}}{\frac{8}{27}} = \frac{37}{216} \times \frac{27}{8} = \frac{37}{64}. \end{aligned}$$

3. An insurance policy pays for a random loss X subject to a deductible of C, where $0 < C < 1$. The loss amount is modeled as a continuous random variable with density function

   $$f(x) = \begin{cases} 2x \text{ for } 0 < x < 1 \\ 0 \qquad\qquad\qquad \text{otherwise} \end{cases}$$

   Given a random loss X, the probability that the insurance payment is less than 0.5 is equal to 0.64.

   (a) Calculate C.
   Solution: Let us first denote by Y the random variable corresponding to the insurance payment from the side of the insurance company. So since there is a deductible C we have

   $$Y = \max(X - C, 0) = \begin{cases} X - C \text{ if } X > C \\ \quad 0 \quad \text{if } X \leq C. \end{cases}$$

We have

$$\begin{aligned} P\left(Y \le \frac{1}{2}\right) &= \frac{64}{100} \iff P\left(X - C \le \frac{1}{2}\right) = \frac{64}{100} \\ &\iff P\left(0 < X \le \frac{1}{2} + C\right) = \frac{64}{100} \\ &\iff \int_0^{\frac{1}{2}+C} 2x dx = \frac{64}{100} \iff \left(\frac{1}{2} + C\right)^2 = \frac{64}{100} \end{aligned}$$

which implies that

$$\frac{1}{2} + C = \pm\sqrt{\frac{64}{100}} = \pm\frac{8}{10}$$

but we have $0 < C < 1$, then $C = \frac{8}{10} - \frac{1}{2} = \frac{3}{10}$.

4. Let X be a continuous random variable with density function

$$f(x) = \begin{cases} \frac{|x|}{10} \text{ for } -2 < x < 4 \\ 0 \qquad\qquad\qquad \text{otherwise} \end{cases}$$

   (a) Calculate the expected value of X.
   Solution: We simply write

$$\begin{aligned} E[X] &= \int_{-2}^{4} x|x|dx = \int_{-2}^{2} x|x|dx + \int_{2}^{4} x^2 dx \\ &= \int_2^4 x^2 dx \quad (\text{since } \int_{-2}^{2} x|x|dx = 0) \\ &= \left[\frac{1}{3}x^3\right]_2^4 = \frac{56}{3}. \end{aligned}$$

5. A device that continuously measures and records seismic activity is placed in a remote region. The time, T, to failure of this device is exponentially distributed with mean 3 years. Since the device will not be monitored during its first two years of service, the time to discovery of its failure is $X = \max(T, 2)$.

   (a) Calculate $E[X]$.
   Solution: We know T is exponentially distributed with parameter $\frac{1}{3}$. So we simply write

$$\begin{aligned} E[X] &= E[\max(T,2)] = 2P(T \le 2) + \int_2^\infty t f_T(t)dt \\ &= 2\int_0^2 f_T(t)dt + \int_2^\infty t f_T(t)dt \\ &= 2\int_0^2 \frac{1}{3}e^{-\frac{1}{3}t}dt + \int_2^\infty t\frac{1}{3}e^{-\frac{1}{3}t}dt \\ &= 2\left[-e^{-\frac{1}{3}t}\right]_0^2 + \left[-e^{-\frac{1}{3}t}(t+3)\right]_2^\infty \\ &= 2 - 2e^{-\frac{2}{3}} + 5e^{-\frac{2}{3}} = 3e^{-\frac{2}{3}} + 2. \end{aligned}$$