PROBABILITY AND STATISTICS - P&P 2

SUBMISSION BY: ELIA YAKIN & ITAY KOREN

Problem 1. Let $X \sim U(1,5)$ (discrete). Define a new random variable $Y = 3^X$

(1) What is the probability mass function of Y?

(2) Compute the expected value of Y.

Answer: (1)

$Supp(Y) = \{y_1, y_2, \dots, y_5\} = \{3, 3^2, 3^3, 3^4, 3^5\}$

Using log rules we can write $Y = 3^X \iff X = log_3 Y = X$, Thus

$$P_Y(Y=y) = P_X(X = log_3 y) \underset{X\sim U}{=} \frac{1}{log_3 y_5 - log_3 y_1 + 1} = \frac{1}{log_3 3^5 - log_3 3^1 + 1} = \frac{1}{5} \ \forall y \in Supp(Y)$$

(2)

$EY = \sum_{y\in Supp(Y)} P_Y(y)\cdot y = \sum_{y\in Supp(Y)} \frac{1}{5}\cdot y = \frac{1}{5}\sum_{y\in Supp(Y)} y = \frac{1}{5}\cdot 363$

Problem 2. According to the British secret intelligence service, during the war, the expected number of bombs ($\mathbb{E}X$) that fall per day in each quarter of London is 2. It is known that the number of bombs that fall in one day in each quarter is a Poisson random variable.

(1) What is the probability that on some day there will be no bombs at all?

(2) What is the probability that at least 4 bombs will fall on some specific quarter in one day?

Answer:

(1) we know that is $X \sim Poiss(\lambda)$ then $EX = \lambda$ hence $X \sim Poiss(2)$ and thus $P_X(0) = \frac{2^0 e^{-2}}{0!} = e^{-2}$

(2) $P(X \geq 4) = 1 - P(X < 3) = 1 - \frac{2^3 e^{-2}}{3!} = 1 - \frac{2^3 e^{-2}}{3\cdot 2} = 1 - \frac{4}{3}\cdot e^{-2}$

Problem 3. Let X be a discrete RV. Show that for any constants $a, b \in \mathbb{R}$:

(1) $E(aX+b) = aEx + b$

(2) $Var(aX+b) = a^2 Var(X)$

Answer:

$$\begin{aligned}
E(aX+b) &\underset{1}{=} \sum_{x\in Supp(X)} (a\cdot x + b)\cdot P(x) \\
&\underset{2}{=} \sum_{x\in Supp(X)} a\cdot x\cdot P(x) + \sum_{x\in Supp(X)} b\cdot P(x) \\
&\underset{3}{=} a\sum_{x\in Supp(X)} x\cdot P(x) + b\sum_{x\in Supp(X)} P(x) \\
&\underset{4}{=} aE(X) + b\cdot 1 = aE(X) + b
\end{aligned}$$

1: Expectation definition; 2: Commutative law over sum; 3: Associative law over sum; 4: from PMF definition $\sum_{x\in Supp(X)} P(x) = 1$

$$\begin{aligned}
Var(aX+b) &\underset{1}{=} E\left((aX+b) - E(aX+b)\right)^2 \\
&\underset{2}{=} E\left(aX+b-(aEX+b)\right)^2 \\
&= E\left(aX - aEX\right)^2 \\
&\underset{3}{=} E\left(a\left(X-EX\right)\right)^2 \\
&\underset{4}{=} a^2 E\left(X-EX\right)^2 \\
&\underset{1}{=} a^2 Var(X)
\end{aligned}$$

1: Variace definition; 2: From (1) 3: Associative law 4: $(A\cdot B)^2 = A^2 \cdot B^2$

Problem 4. Give a simple example of some random variable X and some function $g(\cdot)$ to show that in general $Eg(X) \neq g(EX)$

Answer: let $X \sim U\left(-\frac{1}{2}, \frac{1}{2}\right)$ hence clearly

$$f_X(x) = \begin{cases} 1 & x \in \left[-\frac{1}{2}, \frac{1}{2}\right] \\ 0 & o.w \end{cases}$$

and

$$EX = \int_{-1/2}^{1/2} x \cdot 1 dx = \frac{x^2}{2} \mid_{\frac{-1}{2}}^{\frac{1}{2}} = \frac{1}{2}\left[\left(\frac{1}{2}\right)^2 - \left(-\frac{1}{2}\right)^2\right] = 0$$

let $g(X) = X^2$ thus

$$Eg(X) = \int_{-1/2}^{1/2} x^2 dx = \frac{x^3}{3} \mid_{-1/2}^{1/2} = \frac{1}{3}\left(\left(\frac{1}{2}\right)^3 - \left(-\frac{1}{2}\right)^3\right) = \frac{1}{3}\cdot\frac{2}{8} = \frac{1}{12}$$

which implies

$$g(EX) = 0^2 \neq \frac{1}{12} = Eg(X)$$

Problem 5. In a multiple-choice exam, there are 10 questions. Each has 4 possible answers (only one correct). Alice didn't prepare for the exam, so she guessed all her answers. Let X denote the number of her correct answers.

(1) What is the distribution of X? Write it's PMF.

(2) To pass the test, a student should get 55%. What is the probability that Alice passed the test?

Answer: (1) given Alice guessed all her answers, we can write each question as a bernulli trial with $p = \frac{1}{4}$ and the test as a Binomial distribution. Thus $X \sim Bin\left(10, \frac{1}{4}\right)$. The PMF is defined as:

$$P_X(x) = \binom{10}{x}\left(\frac{1}{4}\right)^x\left(\frac{3}{4}\right)^{10-x}$$

(2)

$$
\begin{aligned}
P(X \geq 6) &\underset{1}{=} 1 - P(X \leq 5) = 1 - \left(\sum_{x=0}^{5} x \cdot P(X = x) \right) \\
&\underset{2}{=} 1 - \sum_{x=0}^{5} P(X = x) \\
&= 1 - \left(\sum_{x=0}^{5} \binom{10}{x} \left(\frac{1}{4}\right)^{x} \left(\frac{3}{4}\right)^{10-x} \right) \\
&= 1 - \left(\sum_{x=0}^{5} \frac{10!}{x!\,(10-x)!} \left(\frac{1}{4}\right)^{x} \left(\frac{3}{4}\right)^{10-x} \right) \\
&\approx 1 - 0.98 = 0.02
\end{aligned}
$$

1: $p = 1 - p^c$ 2: Notice $P(X \geq 5) = P(X = 5) + P(X \geq 4) = P(X = 5) + \cdots + P(X = 0)$

Problem 6. Let X be a random variable with the following density function,

$$
f_X(x) = \begin{cases} ax & x \in (0,1) \\ a & x \in [1,2) \\ a(3-x) & x \in [2,3) \\ 0 & o.w \end{cases}
$$

(1) Find the constant a for which f_X is a density function.
(2) Compute the expectation and variance of X.
(3) Find the cumulative distribution function of X.

Answer: (1) It is not trivial to show that $\int_{(a,b)} z\,dz = \int_{[a,b]} z\,dz$ yet it is intuitive and I assume that it can be done.

$$
\begin{aligned}
1 = \int_{\mathbb{R}} f_X(x)\,dx &= \int_{-\infty}^{0} f_X(x)\,dx + \int_{0}^{1} f_X(x)\,dx \\
&\quad + \int_{1}^{2} f_X(x)\,dx + \int_{2}^{3} f_X(x)\,dx + \int_{3}^{\infty} f_X(x)\,dx \\
&\underset{1}{=} 0 + \int_{0}^{1} f_X(x)\,dx + \int_{1}^{2} f_X(x)\,dx + \int_{2}^{3} f_X(x)\,dx + 0 \\
&= \int_{0}^{1} ax\,dx + \int_{1}^{2} a\,dx + \int_{2}^{3} a(3-x)\,dx \\
&= \frac{a}{2}x^2 \mid_0^1 + ax \mid_1^2 + a\left(3x - \frac{x^2}{2}\right) \mid_2^3 \\
&= \frac{a}{2} + a + a\left(\left(9 - \frac{9}{2}\right) - \left(6 - \frac{4}{2}\right)\right) = \frac{a}{2} + a + a(3 - 2.5) \\
\iff a = \frac{1}{2}
\end{aligned}
$$

(2) given $a=\frac{1}{2}$ we get

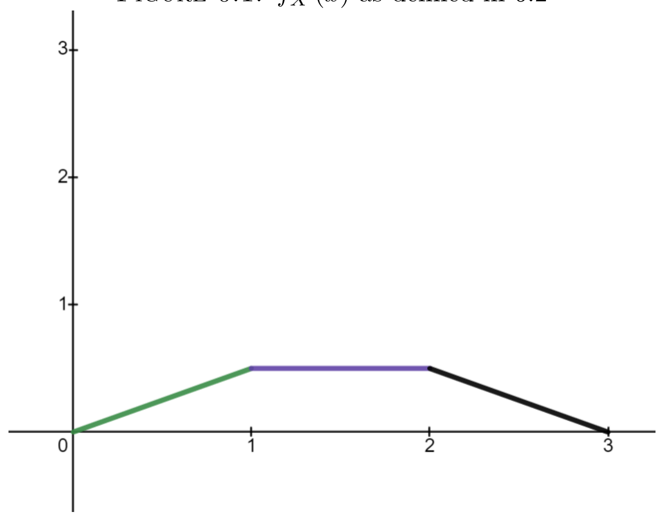
$$f_X(x)=\begin{cases}\frac{x}{2} & x\in(0,1)\\ \frac{1}{2} & x\in[1,2)\\ \frac{3-x}{2} & x\in[2,3)\\ 0 & o.w\end{cases}$$

$$\begin{aligned}EX=\int_{\mathbb{R}}x\cdot f_X(x)\,dx&=\int_0^1\frac{x^2}{2}\,dx+\int_1^2\frac{x}{2}\,dx+\int_2^3\frac{3x-x^2}{2}\,dx\\&=\frac{1}{2}\left[\int_0^1 x^2\,dx+\int_1^2 x\,dx+\int_2^3 3x-x^2\,dx\right]\\&=\frac{1}{2}\left[\frac{1}{3}+\left(\frac{4}{2}-\frac{1}{2}\right)+\left(\left(\frac{27}{2}-\frac{27}{3}\right)-\left(\frac{12}{2}-\frac{8}{3}\right)\right)\right]\\&=\frac{1}{2}\left[\frac{1}{3}+\frac{3}{2}+\frac{7}{6}\right]=\frac{1}{2}\left[\frac{2}{6}+\frac{9}{6}+\frac{7}{6}\right]\\&=\frac{3}{2}\end{aligned}$$

$$\begin{aligned}Var(X)&=E(X-EX)^2\\&=EX^2-E^2X\\&=\int_{\mathbb{R}}x^2\cdot f_X(x)\,dx-\left(\frac{3}{2}\right)^2\\&=\left[\int_0^1\frac{x^3}{2}\,dx+\int_1^2\frac{x^2}{2}\,dx+\int_2^3\frac{3x^2-x^3}{2}\,dx\right]-\frac{9}{4}\\&=\frac{1}{2}\left[\frac{x^4}{4}\mid_0^1+\frac{x^3}{3}\mid_1^2+\left(3\frac{x^3}{3}-\frac{x^4}{4}\right)\mid_2^3\right]-\frac{9}{4}\\&=\frac{1}{2}\left[\frac{1}{4}+\frac{7}{3}+\left(27-\frac{81}{4}-\left(8-\frac{16}{4}\right)\right)\right]-\frac{9}{4}\\&=\frac{1}{2}\left[\frac{3}{12}+\frac{28}{12}+\frac{3}{3}\left(\frac{92}{4}-\frac{81}{4}\right)\right]-\frac{9}{4}\\&=\frac{1}{2}\left[\frac{31}{12}+\frac{33}{12}\right]-\frac{27}{12}=\frac{1}{2}\cdot\frac{64}{12}-\frac{27}{12}\\&=\frac{5}{12}\end{aligned}$$

(3) Reminder $F_X(t):=\int_{-\infty}^t f_X(t)\,dt$.
drawing out $f_X(X)$ we can see it looks like a trapezoid

FIGURE 0.1. $f_X(x)$ as defined in 6.2



Thus we can easily write the CDF as an area of the triangle $S_{[0,1]}(x)$ the rectangle $S_{[1,2]}(x)$ and the trapezoid $S_{[2,3]}(x)$:

$$F_X(x) = \begin{cases} 0 & x \le 0 \\ S_{[0,1]}(x) & x \in (0,1) \\ S_{[1,2]}(x) + S_{[0,1]}(1) & x \in [1,2) \\ S_{[2,3]}(x) + S_{[1,2]}(2) + S_{[0,1]}(1) & x \in [2,3) \\ 1 & x \ge 3 \end{cases} = \begin{cases} 0 & x \le 0 \\ \frac{x \cdot x/2}{2} & x \in (0,1) \\ \frac{(x-1)}{2} \cdot (x-1) + \frac{1}{4} & x \in [1,2) \\ \frac{(4-x)(x-2)}{2} + \frac{3}{4} & x \in [2,3) \\ 1 & x \ge 3 \end{cases}$$

Just to show that it is equal to integration, I show the following (reminder $S_{trapezoid} = \frac{(a+b)\cdot h}{2}$):

$$\begin{aligned} S_{[2,3]}(x) &= \left(\frac{1}{2} + \frac{3-x}{2}\right)(x-2)/2 = \frac{(4-x)(x-2)}{2} \\ &= \frac{4x - 8 - x^2 + 2x}{2} = \frac{6x - x^2 - 8}{2} \\ &= \frac{6x - x^2}{4} - 2 = \int_2^x \frac{3-t}{2}\, dt \end{aligned}$$

Problem 7. Let $X_1, X_2, ...$ be a sequence of independent and identically distributed random variables with distribution $Bin(48, 1/4)$. Let $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$. Use the central limit theorem to calculate $P\left(\overline{X}_{144} > 12.75\right)$

Answer:

$$\begin{aligned}
\mathrm{P}\left(\overline{X}_{144} > 12.75\right) &= 1 - P\left(\overline{X}_{144} \leq 12.75\right) \\
&= 1 - P\left(\sqrt{144}\cdot\frac{\overline{X}_{144} - \mu_X}{\sigma_X} \leq \sqrt{144}\cdot\frac{12.75 - \mu_X}{\sigma_X}\right) \\
&\underset{CLT}{=} 1 - P\left(z \leq 12\cdot\frac{12.75 - \mu_X}{\sigma_X}\right) \\
&= 1 - \phi\left(12\cdot\frac{12.75 - \mu_X}{\sigma_X}\right)
\end{aligned}$$

We need to find μ_X, σ_X.
Notice that for $X \sim Bin\left(n, p\right)$

$$\begin{aligned}
\mu_X = EX = E\left[\sum_{i=1}^{n} X_i\right] &\underset{linearity}{=} \sum_{i=1}^{n} EX_i \underset{iid}{=} nEX_i \\
&= n\left[0\cdot p^0\left(1-p\right)^1 + 1\cdot p^1\left(1-p\right)^0\right] = np \\
\Longrightarrow \mu_X &= 48\cdot\frac{1}{4} = 12 \\
Var\left(X\right) = Var\left(\sum_{i=1}^{n} X_i\right) &\underset{*}{=} \sum_{i=1}^{n} Var\left(X_i\right) \underset{iid}{=} n\cdot Var\left(X_i\right) \\
&= n\cdot\left[E\left[X_i - EX_i\right]^2\right] = n\cdot\left[EX_i^2 - p^2\right] \\
&= n\cdot\left[0^2\cdot p^0\left(1-p\right)^1 + 1^2\cdot p^1\left(1-p\right)^0 - p^2\right] \\
&= n\left[p - p^2\right] = np\left(1-p\right) \\
\Longrightarrow \sigma_X^2 &= 48\cdot\frac{1}{4}\cdot\frac{3}{4} = 9 \implies \sigma_X = 3
\end{aligned}$$

*: $Cov\left(X_i, X_j\right) = 0 \; \forall i \neq j$. Thus

$$\begin{aligned}
P\left(\overline{X}_{144} > 12.75\right) &= 1 - \phi\left(12\cdot\frac{12.75 - 12}{3}\right) \\
&= 1 - \phi\left(12\cdot\frac{0.75}{3}\right) = 1 - \phi\left(3\right) \\
&= 1 - 0.9987 = 0.0013
\end{aligned}$$