HW

2 Probability Home Assignment.

Work has been done by: Ryabykin Aleksey

Problem 1. (Poisson approximation to Binomial)

A group of n people play "Secret Santa". They do it as follows: each puts their name on a slip of paper in a hat, then picks a name randomly from hat without replacement, and then buys a gift for that person. Unfortunately, they overlook the possibility of drawing one's own name. Assume $n \geq 2$. Find the expected value of the number X of people who pick their own names. What is the approximate distribution of X if n is large (specify distribution and parameter)?

Solution: Let's define indicator random variable I_j that indicates whether j-th person has picked his or her own name or not. It is quite clear that $\forall j \in \{1, \dots, n\}$ $P(I_j = 1) = \frac{1}{n} = \mathbb{E}I_j$. After that we can apply the knowledge of expectation linearity and obtain the desired expected value:

$$\mathbb{E}\left[\sum_j I_j\right] = \sum_j \mathbb{E}I_j = \sum_j P(I_j = 1) = n \cdot \frac{1}{n} = 1.$$

By the poisson paradigm, thus for very large n we can approximate x with $y \sim \mathrm{Poiss}(\mathbb{E}x) = \mathrm{Poiss}(1)$:

$$P(x = 0) = \frac{1^0}{0!} \cdot e^{-1} = \frac{1}{e}$$

which is the proability of a $\mathrm{Poiss}(1)$ distribution. ∎

Problem 2. (Poisson process)

Let $X_1, X_2, \dots \sim \mathrm{Be}(p)$ be a series of indepenent random variables. Let $N \sim \mathrm{Poiss}(\lambda)$. Find the PMF of $Y = \sum_{i=1}^{N} X_i$.

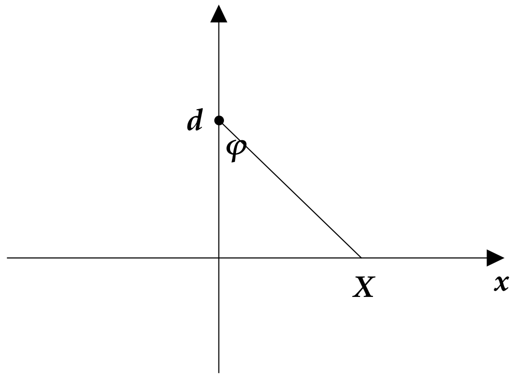
Solution: We have:

$$P(Y = k) = \sum_{n=k}^{\infty} \frac{\lambda^n e^{-\lambda}}{n!} C_n^k p^k (1-p)^{n-k} = \lambda^k e^{-\lambda} \sum_{n=k}^{\infty} \frac{\lambda^{n-k}}{k!(n-k)!} p^k (1-p)^{n-k} =$$
$$= \lambda^k p^k e^{-\lambda} \frac{1}{k!} \sum_{n=0}^{\infty} \frac{\lambda^n (1-p)^n}{n!} = \frac{(\lambda p)^k}{k!} e^{-\lambda} e^{\lambda(1-p)} =$$
$$= \frac{(\lambda p)^k}{k!} e^{-\lambda p} \Longrightarrow$$

$\Rightarrow Y \sim \mathrm{Poiss}(\lambda p)$ ∎

Problem 3. (Cauchy distribution)

Consider $2D$ real plane $\mathbb{R}^2$. Consider a point with coordinates $(0, d)$. Select an angle φ uniformly distributed in $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$. Issue a ray from $(0, d)$ at the angle φ to the y-axis. Denote the point where it intersects x-axis as $(X, 0)$. The distribution of X is called the Caushy distribution. Find its CDF and PDF. Explain, why it does not have an expected value.



Solution: We can find X in terms of angle φ in the following way:

$$\tan\varphi = \frac{X}{d} \Rightarrow X = d\tan\varphi.$$

Then:

$$f(X) = f_\varphi\left(\arctan\left(\frac{X}{d}\right)\right)|J|,$$

where $|J| = \frac{d}{dX}\arctan\frac{X}{d} = \frac{1}{d}\cdot\frac{1}{1+\left(\frac{X}{d}\right)^2} = \frac{d}{d^2+X^2}$. Thus,

$$f\left(\frac{X}{d}\right) = \frac{1}{\pi}\frac{d}{d^2+X^2} = \frac{1}{\pi d\left(1+\left(\frac{X}{d}\right)^2\right)}$$

And CDF:

$$F(x) = \int\limits_{-\infty}^{x}\frac{1}{\pi}\frac{d}{d^2+X^2}dX = \frac{1}{\pi}\arctan\left(\frac{x}{d}\right)$$

$$\mathbb{E}x = \int\limits_{-\infty}^{\infty}\frac{x d dx}{d^2+x^2} = \frac{d}{2}\int\limits_{-\infty}^{\infty}\frac{d(x^2+d^2)}{d^2+x^2} = \frac{d}{2}\ln|d^2+x^2|\Big|_{-\infty}^{\infty} = \infty.$$

■

Problem 4.

Nikita waits $X \sim \mathrm{Gamma}(a, \lambda)$ minutes for the bus to work, and then waits $Y \sim \mathrm{Gamma}(b, \lambda)$ minutes for the bus going home, with X and Y independent. Is the ratio $\frac{X}{Y}$ is independent of the total wait time $X + Y$.

Solution: As we have proved in the classes, that random variables $T = X + Y$ and $W = \frac{X}{X+Y}$ are independent,so any function of W is independent of any function of T. And we have that $\frac{X}{Y}$ is a function of W, since:

$$\frac{X}{Y} = \frac{\frac{X}{X+Y}}{\frac{Y}{X+Y}} = \frac{W}{1-W}$$

So $\frac{X}{Y}$ is independent of $X + Y$. ■

Problem 5. Two subproblems below:

1. Show that if $\mathbb{E}(Y|X) = c$ – is a constant, then X and Y are uncorrelated.
2. Show by example that it is possible to have uncorrelated X and Y such that $\mathbb{E}(Y|X)$ is not a constant.

Solution:

1. Let's show it by using the Adam's law:

> **Theorem: (Adam's law)**
>
> For any r.v.-s X, Y:
> $$\mathbb{E}\left[\mathbb{E}[Y|X]\right] = \mathbb{E}Y$$

Thus we have:

$$\mathbb{E}\left[\mathbb{E}[Y|X]\right] = \mathbb{E}Y = \mathbb{E}c = c.$$

Also

$$\mathbb{E}\left[XY\right] = \mathbb{E}\left[\mathbb{E}\left[XY|X\right]\right] = \mathbb{E}Xc.$$

Now we can compute covariance and correlation furthermore:

$$\operatorname{cov}(X, Y) = \mathbb{E}\left[XY\right] - \mathbb{E}X\mathbb{E}Y = \mathbb{E}Xc - \mathbb{E}Xc = 0$$

We can conclude that correlation is equal to zero either, so these two r.v.-s are uncorrelated.

2. Let's take $X \sim \mathcal{N}(0, 1)$ and $Y = X^2$. As we can see:

$$\operatorname{cov}(X, Y) = \mathbb{E}XY - \mathbb{E}X\mathbb{E}Y = \mathbb{E}X^3 - 0 = 0.$$

the proposed random variables are uncorrelated, but the conditional expectation:

$$\mathbb{E}(Y|X) = \mathbb{E}(X^2|X) = X^2 \neq \text{ const.}$$

is not a constant one.

∎

Problem 6. Conditional variance

Show that $\mathbb{E}\left[\left(Y - \mathbb{E}\left[Y|X\right]\right)^2|X\right] = \mathbb{E}\left[Y^2|X\right] - \left(\mathbb{E}\left[Y|X\right]\right)^2$, so these two expressions for $\operatorname{Var}(Y|X)$ agree.

Solution: Let's perform some computations:

$$\begin{aligned}
\mathbb{E}\left[\left(Y - \mathbb{E}\left[Y|X\right]\right)^2|X\right] &= \mathbb{E}\left[\left(Y^2 - 2Y\mathbb{E}[Y|X] + \mathbb{E}\left[Y|X\right]^2\right)|X\right] = \\
&= \mathbb{E}\left[Y^2|X\right] - 2\mathbb{E}\left[Y\mathbb{E}\left[Y|X\right]|X\right] + \mathbb{E}\left[\mathbb{E}\left[Y|X\right]^2|X\right] = \\
&= \mathbb{E}\left[Y^2|X\right] - 2\mathbb{E}\left[Y|X\right]\mathbb{E}\left[Y|X\right] + \mathbb{E}\left[Y|X\right]^2\mathbb{E}\left[1|X\right] = \\
&= \mathbb{E}\left[Y^2|X\right] - 2\mathbb{E}\left[Y|X\right]^2 + \mathbb{E}\left[Y|X\right]^2 = \\
&= \mathbb{E}\left[Y^2|X\right] - \left(\mathbb{E}\left[Y|X\right]\right)^2
\end{aligned}$$

∎

Problem 7.

Let X be a random variable with mean μ and variance σ^2. Show that:

$$\mathbb{E}(X-\mu)^4 \geq \sigma^4,$$

and use this to show that the kurtosis of X is at least -2.

Solution: It can be shown in the following manner:

$$\mathbb{E}[X-\mu]^4 - \mathbb{E}^2[X-\mu]^2 = \mathrm{Var}^2[X-\mu] \geq 0$$

As it had been shown:

$$\kappa_4 = \frac{\mathbb{E}(X-\mu)^4}{\sigma^4} - 3 \geq \frac{\sigma^4}{\sigma^4} - 3 = -2.$$

■

Problem 8.

Let X be a discrete random variable whose distinct possible values are $x_0, x_1, \dots$, and let $p_k = P(X = x_k)$. The entropy of X is $H(X) = -\sum_k p_k \log p_k$.

1. Find $H(X)$ for $X \sim \mathrm{Geom}(p)$.
2. Let X and Y be i.i.d. discrete random variables. Show that $P(X = Y) \geq 2^{-H(X)}$.

Solution:

1. The entropy can be rewritten in the following terms. Let's define $q = 1 - p$:

$$H(x) = -\sum_k (pq^k)\log\left(pq^k\right) = -\log(p)\sum_k pq^k - \log(q)\sum_k kpq^k =$$

$$= -\log p - \frac{q}{p}\log q$$

since the first series is the sum of a $\mathrm{Geom}(p)$ PMF and the second one is the expected value of a $\mathrm{Geom}(p)$ random variable.

2. Let W be a random variable taking value p_k with probability p_k. By Jensen, $\mathbb{E}(\log(W)) \leq \log(\mathbb{E}W)$, but:

$$\mathbb{E}(\log(W)) = \sum_k p_k \log(p_k) = -H(X).$$

$$\mathbb{E}W = P(X = y),$$

so $-H(X) \leq log P(X = Y)$. Let's take logarithm base 2, thus $P(X = Y) \geq 2^{-H(X)}$.

■

Problem 9.

Let $X_0, X_1, X_2, \dots$ be a Markov chain. Show that X_0, X_2, X_4 is also a Markov chain, and explain why this makes sense intuitively.

Solution: Let's go further and prove it for subchain X_{2n} to make it more intuitive. We can define the present time as $2n$, then we know that $X_{2n+1}, X_{2n+2}, \dots$ is conditionally independent of past $X_0, X_1, \dots, X_{2n-1}$. Thus:

$$P\left(X_{2n+1} = x \middle| X_0 = x_0, \dots X_{2n} = x_{2n}\right) = P\left(X_{2n+1} = x | X_{2n} = x_{2n}\right)$$

■