Probability Theory Homework 4

Gregory Matsnev

November 24, 2025

1 Problem 1

Problem. Find the variance of $X \sim \mathrm{Bin}(n, p)$ using indicator random variables I_j such that $X = I_1 + I_2 + \cdots + I_n$.

Solution. Let I_j be the indicator of success on the j-th trial. Then $I_j \sim \mathrm{Bern}(p)$, hence $\mathbb{E}I_j = p$ and $\mathrm{Var}(I_j) = p(1-p)$. Because the Bernoulli trials are independent,

$$X = \sum_{j=1}^{n} I_j, \qquad \mathbb{E}X = \sum_{j=1}^{n} \mathbb{E}I_j = np,$$

and the independence makes all covariances vanish, so

$$\mathrm{Var}(X) = \sum_{j=1}^{n} \mathrm{Var}(I_j) = np(1-p).$$

2 Problem 2

Problem. Derive the Poisson expectation and variance from its PMF.

Solution. The PMF is $p_k = \mathbb{P}(X = k) = e^{-\lambda}\lambda^k/k!$ for $k = 0, 1, \ldots$. Then

$$\mathbb{E}X = \sum_{k=0}^{\infty} k\, p_k = e^{-\lambda} \sum_{k=1}^{\infty} k \frac{\lambda^k}{k!} = e^{-\lambda}\lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = e^{-\lambda}\lambda e^{\lambda} = \lambda.$$

To compute $\mathbb{E}X^2$ directly from the PMF, expand $k^2 = k(k-1) + k$:

$$\mathbb{E}X^2 = \sum_{k=0}^{\infty} k^2 p_k = \sum_{k=0}^{\infty} k(k-1)p_k + \sum_{k=0}^{\infty} kp_k.$$

The second sum is $\mathbb{E}X = \lambda$ from above. For the first sum, shift the index:

$$\sum_{k=0}^{\infty} k(k-1)e^{-\lambda}\frac{\lambda^k}{k!} = e^{-\lambda}\lambda^2 \sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!} = e^{-\lambda}\lambda^2 e^{\lambda} = \lambda^2.$$

Thus $\mathbb{E}X^2 = \lambda^2 + \lambda$, and by the usual dispersion formula

$$\mathrm{Var}(X) = \mathbb{E}X^2 - (\mathbb{E}X)^2 = (\lambda^2 + \lambda) - \lambda^2 = \lambda.$$

3 Problem 3

Problem. Find the mode, median, and expected value of a random variable X with PDF $\varphi(x) = 3x^2$, where $x \in [0, 1]$. Note: mode, $Mo(X)$, is the "most probable" (in some sense) value of X, i.e. the maximum of PMF/PDF. The median, $Me(X)$, is $F^{-1}(1/2)$ for X with CDF F.

Solution. The density is increasing on $[0, 1]$, so the maximum occurs at the right endpoint and $\mathrm{Mo}(X) = 1$. The CDF is

$$F(x) = \int_0^x 3t^2 \,\mathrm{d}t = x^3, \qquad 0 \le x \le 1.$$

Setting $F(m) = 1/2$ yields the median $m = 2^{-1/3} \approx 0.7937$. The expectation is

$$\mathbb{E}X = \int_0^1 x \cdot 3x^2 \,\mathrm{d}x = 3 \int_0^1 x^3 \,\mathrm{d}x = \frac{3}{4}.$$

4 Problem 4

Problem. Assume that the device repair time is a random variable $X \sim \mathrm{Expo}(\lambda)$. Find the probability that the device repair will take at least 20 days if the average device repair time is 15 days.

Solution. For an exponential distribution, $\mathbb{E}X = 1/\lambda = 15$, so $\lambda = 1/15$. The tail probability is

$$\mathbb{P}(X \ge 20) = e^{-\lambda \cdot 20} = \exp\left(-\frac{20}{15}\right) \approx 0.2636,$$

5 Problem 5

Problem. Consider the Negative Hypergeometric distribution with parameters w, b, and r: an urn contains w white balls and b black balls, which are randomly drawn one by one without replacement, until r white balls have been obtained. Assuming $r \le w$, we say that the number of black balls drawn before drawing the r-th white ball $X \sim \mathrm{NHGeom}(w, b, r)$. Find the $E(X)$ using indicator random variables.

Solution. Index the black balls by $1, \ldots, b$, and let B_i be 1 if the i-th black ball is drawn before the r-th white ball and 0 otherwise. Then $X = \sum_{i=1}^{b} B_i$, so $E(X) = \sum_{i=1}^{b} \mathbb{E}B_i$. In a random permutation of all balls, the relative order of the i-th black ball among the w white balls is uniform over the $w+1$ possible slots (before the first white, between whites, or after the last white). The event $B_i = 1$ occurs precisely when fewer than r white balls appear before this black ball, i.e. the slot number is $0, 1, \ldots, r-1$. Therefore,

$$\mathbb{P}(B_i = 1) = \frac{r}{w+1}, \qquad \mathbb{E}X = \sum_{i=1}^{b} \frac{r}{w+1} = \frac{br}{w+1}.$$

6 Problem 6

Problem. Suppose that Bernoulli trials are being performed in continuous time; i.e. the trials take place at points on a timeline. Assume that the trials are at regularly spaced times $0, \Delta t, 2\Delta t, \ldots$, where t is a small positive number. Let the probability of success of each trial be $\lambda\Delta t$, where λ is a positive constant. Let G be the number of failures before the first success in discrete time, and T be the time of the first success (in continuous time).

(a) Relate G to T and find the CDF of T.

(b) Show that as $\Delta t \to 0$, the CDF of T converges to the Expo(λ) CDF, evaluating all the CDFs at a fixed $t \geq 0$.

Solution. Each discrete-time trial succeeds with probability $\lambda\Delta t$ and fails with probability $1 - \lambda\Delta t$. Thus

$$\mathbb{P}(G = g) = (1 - \lambda\Delta t)^g(\lambda\Delta t), \qquad g = 0, 1, \ldots$$

The first success occurs at time $T = G\Delta t$ because after g failures we wait g spacings of length Δt. For any fixed $t \geq 0$ let $m = \lfloor t/\Delta t \rfloor$. Then

$$F_T(t) = \mathbb{P}(T \leq t) = \mathbb{P}(G \leq m) = 1 - \mathbb{P}(G \geq m + 1) = 1 - (1 - \lambda\Delta t)^{m+1}.$$

As $\Delta t \to 0$, we have $m \sim t/\Delta t$ and the familiar limit $\lim_{x\to\infty}(1 + \frac{y}{x})^x = e^y$ gives

$$\lim_{\Delta t\to 0} F_T(t) = 1 - \lim_{\Delta t\to 0}(1 - \lambda\Delta t)^{(t/\Delta t)+o(1)} = 1 - e^{-\lambda t},$$

which is exactly the Expo(λ) CDF.