CS109 Lecture 7 Worksheet: ANSWER KEY

Variance and the Poisson

Summer 2026

Lecture 7 Worksheet: Answer Key

Instructor / TA copy. Full worked solutions follow.

Problem 1 (Review): Expectation

(a) $X \sim \text{Geo}(0.2)$ (trials until the first success).

(b) $E[X] = 1/p = 1/0.2 = 5$.

(c) $Y \sim \text{Bin}(500, 0.53)$, so $E[Y] = np = 500 \cdot 0.53 = 265$.

Problem 2: Defining Variance

(a) Variance measures the average squared distance of X from its mean μ, i.e. how spread out the distribution is. We square the deviation because $E[X - \mu] = 0$ always (positive and negative deviations cancel); squaring removes the sign so that spread on either side contributes positively.

(b) $\text{Var}(X) = E[X^2] - (E[X])^2$. Variance has **squared** units (e.g. points^2), which is hard to interpret; the standard deviation $\text{SD}(X) = \sqrt{\text{Var}(X)}$ is back in the original units of X (e.g. points), so it is directly comparable to the mean.

Problem 3: Variance of a Die Roll

(a) $\text{Var}(X) = E[X^2] - (E[X])^2 = \frac{91}{6} - (3.5)^2 = 15.1\overline{6} - 12.25 = \frac{35}{12} \approx 2.917$.

(b) $\text{SD}(X) = \sqrt{2.917} \approx 1.708$.

Problem 4: Variance of the Classic Random Variables

(a) Single free throw: $\text{Var} = p(1-p) = 0.53 \cdot 0.47 = 0.2491$.

(b) $Y \sim \text{Bin}(500, 0.53)$: $\text{Var}(Y) = np(1-p) = 500 \cdot 0.53 \cdot 0.47 = 124.55$. Then $\text{SD}(Y) = \sqrt{124.55} \approx 11.16$.

(c) $p(1-p)$ is maximized at $p = \frac{1}{2}$ (its vertex), giving variance $n/4$. A fair coin is the "most unpredictable": with p near 0 or 1 the outcome is nearly certain (little spread), while $p = \frac{1}{2}$ gives the most uncertainty about each trial.

Problem 5: Meet the Poisson

$X \sim \text{Poi}(2.79)$.

(a)
$$P(X=3) = \frac{e^{-2.79}(2.79)^3}{3!} \approx 0.222.$$

(b)
$$P(X=0) = \frac{e^{-2.79}(2.79)^0}{0!} = e^{-2.79} \approx 0.0615.$$

(c) For a Poisson, $E[X] = \text{Var}(X) = \lambda = 2.79$. (Mean and variance are equal, a hallmark of the Poisson.)

Problem 6: Poisson Approximation to a Binomial (pset2: hash_table_poi)

(pset problem)

Problem 7: Web Server Load

$X \sim \text{Poi}(2)$.

(a)
$$P(X<5) = P(X \leq 4) = \sum_{k=0}^{4} \frac{e^{-2}2^k}{k!}.$$

(b) Evaluating: $e^{-2}\left(1+2+2+\frac{8}{6}+\frac{16}{24}\right) = e^{-2}(7.0\overline{6}) \approx 0.947$.

(c) Over 5 seconds the expected count scales with the window: $\lambda = 2 \times 5 = 10$. The Poisson rate is "events per interval," so the interval used for λ must be the same interval the question asks about, or the numbers will not match.

Challenge Problem: DNA Data Storage

(a) Exactly $X \sim \text{Bin}(10^4, 10^{-6})$; since n is huge and p tiny, $X \approx \text{Poi}(\lambda)$ with
$$\lambda = np = 10^4 \cdot 10^{-6} = 0.01.$$

(b)
$$P(X \geq 1) = 1 - P(X=0) = 1 - e^{-0.01} \approx 0.00995.$$

The Poisson approximation is convenient because the exact binomial involves $\binom{10^4}{k}$ and powers like $(10^{-6})^k$, which are awkward and numerically delicate; the Poisson reduces everything to the single parameter $\lambda = 0.01$. (The exact binomial gives 0.00995017, essentially identical.)