Discussion 10B Recap

Tyler Zhu

April 10, 2020

1 Definitions

We'll need the following definitions for today's discussion.

Definition 1 (Expectation). The expectation of a random variable X is

$$\mathbb{E}[X] = \sum_{x \in \mathcal{A}} x\mathbb{P}(X = x).$$

Definition 2 (Variance). The variance of a random variable X is

$$\mathrm{Var}(X) = \sum_{x \in \mathcal{A}} (x - \mathbb{E}[X])^2 \mathbb{P}(X = x) = \mathbb{E}[X^2] - \mathbb{E}[X]^2.$$

Remember that expectation of a function of any random variable is an average of the values of the r.v. with the function applied, weighted by the probabilities, as one would expect. Formally,

$$\mathbb{E}[f(X)] = \sum_{x \in \mathcal{A}} f(x)\mathbb{P}(X = x).$$

This is sometimes referred to as the *law of the unconscious statistician.*

2 Poisson Distributions

Poisson distributions model rare events, such as the number of cases of disease, number of births per hour, or number of arrivals of a bus in an hour. They are defined in terms of a parameter λ, referred to as the *intensity* or the *rate*, which specifies the average number of times an event occurs in a continuous interval.

Definition 3 (Poisson Distribution). A r.v. X is a *Poisson Distribution* with parameter λ if it has probability mass function

$$\mathbb{P}(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}$$

for $k = 0, 1, \ldots$ (i.e. X is a nonnegative random variable). We write $X \sim \mathrm{Poisson}(\lambda)$ to denote this.

There are a number of implications here:

- Since this is a distribution, $\sum_{k=0}^{\infty} \frac{e^{-\lambda}\lambda^k}{k!} = 1$.
- $\mathbb{E}[X] = \mathrm{Var}(X) = \lambda$, as one would expect since λ is the average number of occurrences.
- Occurences that occur in disjoint subintervals or subregions are independent.
- Occurences on some subinterval A of the entire interval I are distributed $\sim \mathrm{Poisson}(\lambda|A|/|I|)$.

Finally, there is this neat property of Poisson random variables that we will use often (Theorem 19.5 in the notes).

Theorem 4. Let $X \sim \text{Poisson}(\lambda)$ and $Y \sim \text{Poisson}(\mu)$ be independent Poisson random variables. Then $X + Y \sim \text{Poisson}(\lambda + \mu)$.

Notice that this doesn't hold for dependent Poisson r.v.'s, and it also doesn't hold for independent geometric r.v.'s.

It is also useful to have an alternate view of these distributions. Since Poisson distributions measure the number of arrivals in a given interval, we could also think of them as a Binomial distribution. To be precise, if $X \sim \text{Poisson}(\lambda)$, then on an interval of length n, we would expect there to be about λ/n arrivals on average. If we take n to be big enough, then we can make the assumption that there is at most one arrival in each chunk of length 1, so this is just distributed as $\sim \text{Bin}(n, \lambda/n)$!

In fact, as $n \to \infty$, if $X \sim \text{Bin}(n, \lambda/n)$, then $\mathbb{P}(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}$, the PMF of a Poisson distribution.

3 Computing Variance of a Sum of Indicators

Suppose I have a random variable $X = X_1 + \cdots + X_n$ where X_i are random indicator variables. Then the expectation is just

$$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X_i]$$

by the linearity of expectation. The variance however is not so easy, but is important to know as it comes up time and time again. We begin by using the formula above to write it as $\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$, so we only need concern ourselves with the first term. Expanding let's us write it as

$$\begin{aligned}
\mathbb{E}[X^2] &= \mathbb{E}\left[(X_1 + \cdots + X_n)^2\right] \\
&= \mathbb{E}\left[\sum_{i=1}^{n} X_i^2 + \sum_{i \neq j} X_i X_j\right] \\
&= \mathbb{E}\left[\sum_{i=1}^{n} X_i\right] + \mathbb{E}\left[\sum_{i \neq j} X_i X_j\right] \\
&= \mathbb{E}[X] + \sum_{i \neq j} \mathbb{E}[X_i X_j].
\end{aligned}$$

We got from the second line to the third line since X_i only takes on the values 0 or 1, so $X_i^2 = X_i$ for all values it takes on. Thus, we can reduce all computations of $\mathbb{E}[X^2]$ to computing $\mathbb{E}[X_i X_j]$, which in term depends only on what $\mathbb{P}(X_i = 1 \cap X_j = 1)$ is.