Math 3070, Applied Statistics

Section 1

September 13, 2019

Lecture Outline, 9/13

Section 3.6

- Poisson Distribution

Poisson Process

Consider a process where events occur at random times, such as

- The arrival times of customers at a store
- Clicks of a Geiger counter exposed to a radioactive material
- Webpage requests on an internet server
- Incoming calls to a customer service center

Poisson Process

Consider a process where events occur at random times, such as

- The arrival times of customers at a store
- Clicks of a Geiger counter exposed to a radioactive material
- Webpage requests on an internet server
- Incoming calls to a customer service center

Such a process is called a *Poisson process* if the following assumptions hold:

1. The mean number of events which occur in a time interval of length t is λt, where λ is a constant, called the *rate* of the Poisson process.
2. Events occur only one at a time.
3. The number of events which occur in a time interval is independent of the number and timing of past events.

Poisson Distribution

Given a Poisson process with rate λ, we want to find the pmf of the number of events X which occur in the time interval $I = [0, t]$.

Poisson Distribution

Given a Poisson process with rate λ, we want to find the pmf of the number of events X which occur in the time interval $I = [0, t]$.

- By assumption (1), the mean of X is $\mu = \lambda t$.
- Divide the interval $[0, t]$ into equal-width subintervals $I_1, \ldots, I_n$, each of length t/n.
- Let X_k be the number of events which occur in the subinterval I_k, so $X = X_1 + X_2 + \cdots + X_n$.
- By assumption (1), $E(X_k) = \lambda \cdot \frac{t}{n} = \frac{\mu}{n}$.
- By assumption (2), if n is large, then the probability of more than one event occuring in any given interval I_k is very small.
- So we may approximate X_k as a Bernoulli random variable with parameter $p = \frac{\mu}{n}$.
- By assumption (3), the random variables $X_1, \ldots, X_n$ are independent. Thus X is approximately binomial, $Bin(n, \frac{\mu}{n})$.

Poisson Distribution

Recall from calculus that

$$\lim_{n\to\infty}\left(1+\frac{r}{n}\right)^n = e^r$$

We use this to find the pmf of a Poisson random variable X. We argued that for large n, X is approximately binomial, $Bin(n, \frac{\mu}{n})$, so

$$\begin{aligned}
P(X=x) &\approx \binom{n}{x}(\mu/n)^x(1-\mu/n)^{n-x} \\
&= \frac{n(n-1)\cdots(n-x+1)}{x!}(\mu/n)^x(1-\mu/n)^{n-x} \\
&= \frac{n(n-1)\cdots(n-x+1)}{nn\cdots n}(1-\mu/n)^n(1-\mu/n)^{-x}\cdot\frac{\mu^x}{x!} \\
&\to 1\cdot e^{-\mu}\cdot 1\cdot\frac{\mu^x}{x!} \qquad\qquad (\text{as } n\to\infty) \\
&= \frac{e^{-\mu}\mu^x}{x!}
\end{aligned}$$

Poisson Distribution

Given a Poisson process with rate λ, the number X of events which occur in a time interval of length t is a **Poisson** random variable with mean $\mu = \lambda t$. $X \sim Pois(\mu)$ The possible values of X are $0, 1, 2, \ldots$

$$P(X = x) = \frac{e^{-\mu}\mu^x}{x!}$$

Poisson Distribution

Given a Poisson process with rate λ, the number X of events which occur in a time interval of length t is a **Poisson** random variable with mean $\mu = \lambda t$. $X \sim Pois(\mu)$ The possible values of X are $0, 1, 2, \ldots$

$$P(X = x) = \frac{e^{-\mu}\mu^x}{x!}$$

Example: Suppose that at a small store, customers arrive at an average rate of 6 per hour. What is the probability that during a given hour only 3 or fewer customers will arrive?

Poisson Distribution

Given a Poisson process with rate λ, the number X of events which occur in a time interval of length t is a **Poisson** random variable with mean $\mu = \lambda t$. $X \sim Pois(\mu)$ The possible values of X are $0, 1, 2, \ldots$

$$P(X = x) = \frac{e^{-\mu}\mu^x}{x!}$$

Example: Suppose that at a small store, customers arrive at an average rate of 6 per hour. What is the probability that during a given hour only 3 or fewer customers will arrive?

Assuming a Poisson process,

$$\begin{aligned} P(X \leq 3) &= P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) \\ &= \frac{e^{-6}6^0}{0!} + \frac{e^{-6}6^1}{1!} + \frac{e^{-6}6^2}{2!} + \frac{e^{-6}6^3}{3!} \\ &= e^{-6}(1 + 6 + 18 + 36) = 61e^{-6} \approx .151 \end{aligned}$$

Poisson Distribution, Sum of PMF

Recall from calculus that

$$e^{\mu} = \sum_{x=0}^{\infty} \frac{\mu^x}{x!}$$

Poisson Distribution, Sum of PMF

Recall from calculus that

$$e^{\mu} = \sum_{x=0}^{\infty} \frac{\mu^x}{x!}$$

Therefore, the pmf $f(x)$ of a Poisson random variable satisfies

$$\begin{aligned}
\sum_{x=0}^{\infty} f(x) &= \sum_{x=0}^{\infty} \frac{e^{-\mu}\mu^x}{x!} \\
&= e^{-\mu} \sum_{x=0}^{\infty} \frac{\mu^x}{x!} \\
&= e^{-\mu} e^{\mu} \\
&= 1
\end{aligned}$$

which shows that $f(x)$ is a valid pmf.

Mean of Poisson Distribution

We may directly calculate the mean of a Poisson random variable X based on the pmf:

Mean of Poisson Distribution

We may directly calculate the mean of a Poisson random variable X based on the pmf:

$$\begin{aligned}
E(X) &= \sum_{x=0}^{\infty} x \cdot \frac{e^{-\mu}\mu^{x}}{x!} \\
&= \sum_{x=1}^{\infty} \frac{e^{-\mu}\mu^{x}}{(x-1)!} \\
&= \sum_{x=0}^{\infty} \frac{e^{-\mu}\mu^{x+1}}{x!} \\
&= \mu \sum_{x=0}^{\infty} \frac{e^{-\mu}\mu^{x}}{x!} \\
&= \mu
\end{aligned}$$

This is what we expected based on the definition.

Variance of Poisson Distribution

We can also calculate the variance:

$$
\begin{aligned}
E(X^2) &= \sum_{x=0}^{\infty} x^2 \cdot \frac{e^{-\mu}\mu^x}{x!} \\
&= \sum_{x=1}^{\infty} x \cdot \frac{e^{-\mu}\mu^x}{(x-1)!} \\
&= \sum_{x=0}^{\infty} (x+1) \frac{e^{-\mu}\mu^{x+1}}{x!} \\
&= \sum_{x=0}^{\infty} x \frac{e^{-\mu}\mu^{x+1}}{x!} + \sum_{x=0}^{\infty} \frac{e^{-\mu}\mu^{x+1}}{x!} \\
&= \sum_{x=1}^{\infty} \frac{e^{-\mu}\mu^{x+1}}{(x-1)!} + \mu \sum_{x=0}^{\infty} \frac{e^{-\mu}\mu^x}{x!} \\
&= \sum_{x=0}^{\infty} \frac{e^{-\mu}\mu^{x+2}}{x!} + \mu = \mu^2 + \mu
\end{aligned}
$$

So $V(X) = E(X^2) - [E(X)]^2 = \mu^2 + \mu - \mu^2 = \mu$.

Example

Suppose that at random times a system suffers breakdowns requiring immediate repairs. If the system breaks down at a rate of once per year, what is the probability that the system will break down 3 or more times in one year?

Example

Suppose that at random times a system suffers breakdowns requiring immediate repairs. If the system breaks down at a rate of once per year, what is the probability that the system will break down 3 or more times in one year?

Solution: The given information suggests that the number X of breakdowns in a year is a Poisson random variable with mean $\mu = \lambda t = 1(1) = 1$.

Example

Suppose that at random times a system suffers breakdowns requiring immediate repairs. If the system breaks down at a rate of once per year, what is the probability that the system will break down 3 or more times in one year?

Solution: The given information suggests that the number X of breakdowns in a year is a Poisson random variable with mean $\mu = \lambda t = 1(1) = 1$.

$$\begin{aligned}
P(X \geq 3) &= 1 - P(X \leq 2) \\
&= 1 - P(X = 0) - P(X = 1) - P(X = 2) \\
&= 1 - \frac{e^{-1}1^0}{0!} - \frac{e^{-1}1^1}{1!} - \frac{e^{-1}1^2}{2!} \\
&= 1 - e^{-1}\left(1 + 1 + \frac{1}{2}\right) \\
&= 1 - \frac{5e^{-1}}{2} \approx .080
\end{aligned}$$

Approximating Binomial for Rare Events

Binomial random variables with small p and large n can also be approximated with a Poisson random variable, $\mu = np$.

Approximating Binomial for Rare Events

Binomial random variables with small p and large n can also be approximated with a Poisson random variable, $\mu = np$.

$$\begin{aligned}
P(X = x) &= \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \\
&= \frac{n(n-1)\cdots(n-x+1)}{x!} \left(\frac{\mu}{n}\right)^x \left(1 - \frac{\mu}{n}\right)^{n-x} \\
&= \frac{1}{x!} \frac{n^x(1-\frac{1}{n})\cdots(1-\frac{x-1}{n})}{n^x} \mu^x \left(1 - \frac{\mu}{n}\right)^{n-x} \\
&\to \frac{e^{-\mu}\mu^x}{x!}, \quad (\text{as } n \to \infty)
\end{aligned}$$

Approximating Binomial for Rare Events

Binomial random variables with small p and large n can also be approximated with a Poisson random variable, $\mu = np$.

$$
\begin{aligned}
P(X = x) &= \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \\
&= \frac{n(n-1)\cdots(n-x+1)}{x!} \left(\frac{\mu}{n}\right)^x \left(1 - \frac{\mu}{n}\right)^{n-x} \\
&= \frac{1}{x!} \frac{n^x (1 - \frac{1}{n}) \cdots (1 - \frac{x-1}{n})}{n^x} \mu^x \left(1 - \frac{\mu}{n}\right)^{n-x} \\
&\to \frac{e^{-\mu} \mu^x}{x!}, \quad (\text{as } n \to \infty)
\end{aligned}
$$

Note: n and p do not appear as parameters of the exponential distribution, but can be related by μ.

Approximating Binomial for Rare Events

Consider $E(X)$ and $V(X)$ of $X \sim bin(n, p)$ as $np \to \mu$ while $n \to \infty$ and $p \to 0$.

$$E(X) = np \to \mu$$

$$V(X) = np(1 - p) \to \mu$$

Parameters match in the limit.

Example, Approximating Binomial for Rare Events

The Prussian army studied the likelihood of a soldier being killed by a horse kick. Over twenty years, it was observed that 122 soldiers from 10 corps were killed horse kick. The size of each corps is 10,000 soldiers. Approximate the probability that no Prussian soldiers from two corps are killed by horse kick in one year.

Example, Approximating Binomial for Rare Events

The Prussian army studied the likelihood of a soldier being killed by a horse kick. Over twenty years, it was observed that 122 soldiers from 10 corps were killed horse kick. The size of each corps is 10,000 soldiers. Approximate the probability that no Prussian soldiers from two corps are killed by horse kick in one year.

Estimate p from the historical data,

$$p \approx \frac{122}{10000 \cdot 20 \cdot 10}.$$

Compute μ.

$$\mu = np = 2 \cdot 10000 \frac{122}{10000 \cdot 20 \cdot 10} = \frac{244}{200} = 1.22$$

Estimate with X with a $exp(1.22)$ random variable.

$$P(X = 0) = \frac{e^{-1.22} 1.22^0}{0!} = e^{-1.22} \approx 0.29523016692$$

Poisson Distribution, Summary

- $X \sim Pois(\mu)$ mean X is a poisson random variable with rate μ and its PMF is
$$P(X = x) = \frac{e^{-\mu}\mu^{x}}{x!}$$
- $$E(X) = V(X) = \mu$$
- When modeling number of events in an interval, $\mu = \lambda t$. λ is the rate of occurance per unit time. Sometimes, μ is call the rate instead of λ. Questions in this class will always assume μ is the mean.
- When approximating the number of rare events in a large population, $\mu = np$.

Midterm September 18, Information

- There is a midterm on September 18th in class. Calculator and notes allowed.
- Study the quizzes, summary slides, and homework. See the Canvas 'Files' tab for information.
- Review on September 16th. Come with questions.
- Reschedule by Wednesday, September 11, if needed. No makeup or late exams.
- One question from this week's material.
- No quiz or homework due on exam weeks. Material is shifted to the later week.