3.22 Poisson Distribution

Poisson Process

Poisson Distribution

Properties

Poisson Process

Let $N(t)$ be a **counting process**. That is, $N(t)$ is the number of occurrences (or arrivals, or events) of some process over the time interval $[0, t]$. $N(t)$ looks like a step function.

Examples: $N(t)$ could be any of the following.

(a) Cars entering a shopping center (time).

(b) Defects on a wire (length).

(c) Raisins in cookie dough (volume).

Let $\lambda > 0$ be the average number of occurrences per unit time (or length or volume).

In the above examples, we might have:

(a) $\lambda = 10/\text{min}$. (b) $\lambda = 0.5/\text{ft}$. (c) $\lambda = 4/\text{in}^3$.

A Poisson process is a specific counting process. . .

First, some notation: $o(h)$ is a generic function that goes to zero faster than h goes to zero.

Definition: A **Poisson process** is one that satisfies the following assumptions:

(1) There is a short enough interval of time, say of length h, such that, for all t,

$$\begin{aligned}
\Pr(N(t+h)-N(t)=0) &= 1-\lambda h+o(h) \\
\Pr(N(t+h)-N(t)=1) &= \lambda h+o(h) \\
\Pr(N(t+h)-N(t)\geq 2) &= o(h)
\end{aligned}$$

(2) If $t_1 < t_2 < t_3 < t_4$, then $N(t_4)-N(t_3)$ and $N(t_2)-N(t_1)$ are *indep* RV's.

English translation of Poisson process assumptions.

(1) Arrivals basically occur one-at-a-time, and then at rate λ/unit time. (We must make sure that λ doesn't change over time.)

(2) The numbers of arrivals in two disjoint time intervals are indep.

Poisson Process Example: Neutrinos hit a detector. Occurrences are rare enough so that they really do happen one-at-a-time. You never get arrivals of groups of neutrinos. Further, the rate doesn't vary over time, and all arrivals are indep of each other.

Anti-Example: Customers arrive at a restaurant. They show up in groups, not one-at-a-time. The rate varies over the day (more at dinnertime). Arrivals may not be indep. This ain't a Poisson process.

Poisson Distribution

Definition: Let X be the number of occurrences in a Poisson(λ) process in a *unit interval* of time. Then X has the **Poisson distribution** with parameter λ.

Notation: $X \sim$ Pois(λ).

Theorem/Definition: $X \sim$ Pois(λ) $\Rightarrow$
Pr($X = k$) $= e^{-\lambda}\lambda^k/k!$, $k = 0, 1, 2, \ldots$.

Remark: The value of λ can be changed simply by changing the units of time.

Example:

$X = \#$ calls to a switchboard in 1 minute $\sim$ Pois(3)

$Y = \#$ calls to a switchboard in 5 minutes $\sim$ Pois(15)

$Z = \#$ calls to a switchboard in 10 sec $\sim$ Pois(0.5)

Properties

Theorem: $X \sim \text{Pois}(\lambda) \Rightarrow$ mgf is $M_X(t) = e^{\lambda(e^t-1)}$.

Proof:

$$\begin{aligned} M_X(t) &= \mathsf{E}[e^{tX}] = \sum_{k=0}^{\infty} e^{tk}\left(\frac{e^{-\lambda}\lambda^k}{k!}\right) \\ &= e^{-\lambda}\sum_{k=0}^{\infty}\frac{(\lambda e^t)^k}{k!} \\ &= e^{-\lambda}e^{\lambda e^t}. \end{aligned}$$

Theorem: $X \sim \text{Pois}(\lambda) \Rightarrow \text{E}[X] = \text{Var}(X) = \lambda$.

Proof (using mgf):

$$\begin{aligned}
\text{E}[X] &= \frac{d}{dt} M_X(t)\bigg|_{t=0} \\
&= \frac{d}{dt} e^{\lambda(e^t - 1)}\bigg|_{t=0} \\
&= \lambda e^t M_X(t)\bigg|_{t=0} \quad \text{(chain rule)} \\
&= \lambda \quad \text{(after algebra)}.
\end{aligned}$$

Similarly,

$$\begin{aligned}
\mathsf{E}[X^2] &= \frac{d^2}{dt^2}M_X(t)\Big|_{t=0} = \frac{d}{dt}\Big(\frac{d}{dt}M_X(t)\Big)\Big|_{t=0} \\
&= \lambda\frac{d}{dt}(e^t M_X(t))\Big|_{t=0} \\
&= \lambda\Big[e^t M_X(t) + e^t\frac{d}{dt}M_X(t)\Big]\Big|_{t=0} \\
&= \lambda e^t\Big[M_X(t) + \lambda e^t M_X(t)\Big]\Big|_{t=0} \\
&= \lambda(1+\lambda).
\end{aligned}$$

Thus,

$$\mathsf{Var}(X) \;=\; \mathsf{E}[X^2] - (\mathsf{E}[X])^2 \;=\; \lambda(1+\lambda) - \lambda^2 \;=\; \lambda.$$

Done.

Example: Calls to a switchboard arrive as a Poisson process with rate 3 calls/min.

Let $X =$ number of calls in 40 sec. So $X \sim$ Pois(2).

$\mathsf{E}[X] = \mathsf{Var}(X) = 2$, $\Pr(X \le 3) = \sum_{k=0}^{3} e^{-2} 2^k / k!$

Theorem (Additive Property of Poissons): Suppose $X_1, \ldots, X_n$ are *indep* with $X_i \sim$ Pois(λ_i), $i = 1, \ldots, n$. Then

$$Y \equiv \sum_{i=1}^{n} X_i \sim \text{Pois}\Big(\sum_{i=1}^{n} \lambda_i\Big).$$

Proof:

$$\begin{aligned} M_Y(t) &= \prod_{i=1}^{n} M_{X_i}(t) \quad (X_i\text{'s indep}) \\ &= \prod_{i=1}^{n} e^{\lambda_i(e^t-1)} = e^{(\sum_{i=1}^{n} \lambda_i)(e^t-1)}, \end{aligned}$$

which is the mgf of the Pois($\sum_{i=1}^{n} \lambda_i$) distribution.