Poisson Distributions

Engineering Statistics
Section 3.6

Josh Engwer

TTU

24 February 2016

Siméon Denis Poisson (1781-1840)

Poisson Random Variables (Applications)

Poisson random variables are used to model the following:

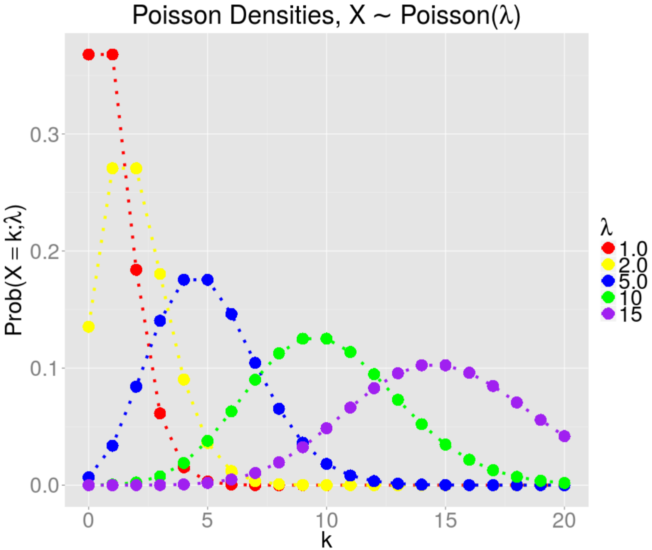
- Number of **arrivals** over a fixed **time period** Δt
    - # radioactive decays of $1\mu g$ of Iodine-123 in 1/1000 second
    - # phone calls a dispatcher received in 45 minutes
    - # emails an account received in two hours
    - # car accidents at a dangerous intersection in four weeks
    - # insurance claims from a given demographic in six months
    - # industrial accidents at a factory in five years
    - # wars started in a continent in three centuries
- Number of **arrivals** over a fixed **length** ΔL
    - # mutations in a strand of DNA
    - # blemishes in a spool of copper wire
- Number of **arrivals** over a fixed **area** ΔA
    - # chocolate chips in a large cookie
- Number of **arrivals** over a fixed **volume** ΔV
    - # yeast cells used in brewing a glass of Guinness beer

Poisson Random Variables (Summary)

Proposition

Notation	$X \sim \textit{Poisson}(\lambda), \quad \lambda > 0$
Parameter(s)	$\lambda = \alpha\Delta t$ *s.t.* $\lambda \equiv$ *Expected # Arrivals over Time Period*; $\alpha \equiv$ *Expected # Arrivals per Unit Time*; $\Delta t \equiv$ *Time period*
Support	$\textit{Supp}(X) = \{0, 1, 2, 3, \cdots\}$
Density (*pmf*)	$p_X(k;\lambda) := \frac{\lambda^k}{k!}e^{-\lambda}$
Mean *Variance*	$\mathbb{E}[X] = \lambda$ $\mathbb{V}[X] = \lambda$
Model(s)	*Number of arrivals over a fixed time period* Δt *Number of arrivals over a fixed space* ΔA *or* ΔV
Assumption(s)	$\mathbb{P}(\textit{No arrivals during time period } \Delta t) \approx 1 - \alpha\Delta t$ $\mathbb{P}(\textit{Exactly one arrival during time period } \Delta t) \approx \alpha\Delta t$ $\mathbb{P}(\textit{More than one arrival during time period } \Delta t) \approx 0$ *# Arrivals during* ***disjoint*** *time periods are* ***independent***

Poisson Density Plots (pmf's)



Remember, the only meaningful values of k for Poisson r.v.'s are $0, 1, 2, 3, \cdots$

Verification that Poisson pmf truly is a valid pmf

It's not immediately obvious that $p_X(k;\lambda) = \frac{\lambda^k}{k!}e^{-\lambda}$ is a pmf, so let's prove it:

- Non-negativity on its support:

  Let $k \in \text{Supp}(X) = \overline{\mathbb{N}} = \{0, 1, 2, 3, \cdots\}$ and $\lambda > 0$.
  Then $k! > 0, \quad \lambda^k > 0,$ and $e^{-\lambda} = \frac{1}{e^\lambda} > 0 \implies p_X(k;\lambda) > 0$

- Universal Summation of Unity:

$$\sum_{k \in \text{Supp}(X)} p_X(k;\lambda) = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!}e^{-\lambda} = e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!} \overset{TAYLOR}{=} e^{-\lambda} \cdot e^{\lambda} = 1$$

Recall from Calculus II that the Taylor Series about $x = 0$ for e^x is $\sum_{k=0}^{\infty} \frac{x^k}{k!}$.

Mean of Poisson(λ) random variable (Proof)

Let random variable $X \sim$ Poisson(λ), where $\lambda > 0$. Then:

$$\mathbb{E}[X] = \sum_{k \in \text{Supp}(X)} k \cdot p_X(k;\lambda) = \sum_{k=0}^{\infty} k \cdot \frac{\lambda^k}{k!} e^{-\lambda} = \sum_{k=1}^{\infty} k \cdot \frac{\lambda^k}{k!} e^{-\lambda}$$

$$= \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} e^{-\lambda} = e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} \overset{CV}{=} e^{-\lambda} \sum_{j=0}^{\infty} \frac{\lambda^{j+1}}{j!}$$

$$= \lambda e^{-\lambda} \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} \overset{TAYLOR}{=} \lambda e^{-\lambda} \cdot e^{\lambda} = \lambda \cdot e^{-\lambda+\lambda} = \lambda \cdot e^0 = \lambda \cdot 1 = \lambda$$

$\therefore \quad \mathbb{E}[X] = \lambda$ QED

CV: Let $j = k - 1 \iff k = j + 1$.

Then
$$\begin{array}{lllll} k = \infty & \implies & j + 1 = \infty & \implies & j = \infty - 1 = \infty \\ k = 1 & \implies & j + 1 = 1 & \implies & j = 1 - 1 = 0 \end{array}$$

Recall from Calculus II that the Taylor Series about $x = 0$ for e^x is $\displaystyle\sum_{k=0}^{\infty} \frac{x^k}{k!}$.

Variance of Poisson(λ) random variable (Proof)

Let random variable $X \sim$ Poisson(λ), where $\lambda > 0$. Then:

$$\mathbb{E}[X^2] = \sum_{k \in \text{Supp}(X)} k^2 \cdot p_X(k;\lambda) = \sum_{k=0}^{\infty} k^2 \cdot \frac{\lambda^k}{k!} e^{-\lambda} = \sum_{k=1}^{\infty} k^2 \cdot \frac{\lambda^k}{k!} e^{-\lambda}$$

$$= e^{-\lambda} \sum_{k=1}^{\infty} k \cdot \frac{\lambda^k}{(k-1)!} = \lambda e^{-\lambda} \sum_{k=1}^{\infty} k \cdot \frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\partial}{\partial \lambda} \left[\frac{\lambda^k}{(k-1)!} \right]$$

$$\overset{(*)}{=} \lambda e^{-\lambda} \frac{d}{d\lambda} \left[\sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} \right] \overset{CV}{=} \lambda e^{-\lambda} \frac{d}{d\lambda} \left[\sum_{j=0}^{\infty} \frac{\lambda^{j+1}}{j!} \right] = \lambda e^{-\lambda} \frac{d}{d\lambda} \left[\lambda \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} \right]$$

$$= \lambda e^{-\lambda} \frac{d}{d\lambda} \left[\lambda e^{\lambda} \right] = \lambda e^{-\lambda} \left[e^{\lambda} + \lambda e^{\lambda} \right] = \lambda + \lambda^2$$

$$\therefore \quad \mathbb{V}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = (\lambda + \lambda^2) - (\lambda)^2 = \lambda \qquad \text{QED}$$

$(*)$ Interchanging summation & differentiation works here, but not in general. Take **Advanced Calculus** for the painful details.

Probabilities of Poisson rv's via Poisson cdf Pois$(x; \lambda)$

It can get quite tedious using the Poisson pmf.
It's easier to use the Poisson cdf, but the cdf has no elementary closed-form!
Instead, tables of numerical values of the Poisson cdf are used instead.

Probabilities of Poisson rv's via Poisson cdf Pois$(x; \lambda)$

	Expected/Average # Arrivals over entire Time Period (λ)										
x	**0.1**	**0.2**	**0.25**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.75**	**0.8**	**0.9**
0	0.90484	0.81873	0.77880	0.74082	0.67032	0.60653	0.54881	0.49659	0.47237	0.44933	0.40657
1	0.99532	0.98248	0.97350	0.96306	0.93845	0.90980	0.87810	0.84420	0.82664	0.80879	0.77248
2	0.99985	0.99885	0.99784	0.99640	0.99207	0.98561	0.97688	0.96586	0.95949	0.95258	0.93714
3	1.00000	0.99994	0.99987	0.99973	0.99922	0.99825	0.99664	0.99425	0.99271	0.99092	0.98654
4	1.00000	1.00000	0.99999	0.99998	0.99994	0.99983	0.99961	0.99921	0.99894	0.99859	0.99766
5	1.00000	1.00000	1.00000	1.00000	1.00000	0.99999	0.99996	0.99991	0.99987	0.99982	0.99966
6	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	0.99999	0.99999	0.99998	0.99996
7	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000

Suppose $X \sim$ Poisson$(\lambda = 0.8)$. Then:

$$\begin{aligned} \mathbb{P}(X \leq 1) &= ?????? \\ \mathbb{P}(X \leq 2) &= ?????? \\ \mathbb{P}(X = 2) &= ?????? \\ \mathbb{P}(X > 2) &= ?????? \end{aligned}$$

Probabilities of Poisson rv's via Poission cdf Pois$(x; \lambda)$

	Expected/Average # Arrivals over entire Time Period (λ)										
x	**0.1**	**0.2**	**0.25**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.75**	**0.8**	**0.9**
0	0.90484	0.81873	0.77880	0.74082	0.67032	0.60653	0.54881	0.49659	0.47237	0.44933	0.40657
1	0.99532	0.98248	0.97350	0.96306	0.93845	0.90980	0.87810	0.84420	0.82664	0.80879	0.77248
2	0.99985	0.99885	0.99784	0.99640	0.99207	0.98561	0.97688	0.96586	0.95949	0.95258	0.93714
3	1.00000	0.99994	0.99987	0.99973	0.99922	0.99825	0.99664	0.99425	0.99271	0.99092	0.98654
4	1.00000	1.00000	0.99999	0.99998	0.99994	0.99983	0.99961	0.99921	0.99894	0.99859	0.99766
5	1.00000	1.00000	1.00000	1.00000	1.00000	0.99999	0.99996	0.99991	0.99987	0.99982	0.99966
6	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	0.99999	0.99999	0.99998	0.99996
7	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000

Suppose $X \sim$ Poisson$(\lambda = 0.8)$. Then:

$$\begin{aligned}
\mathbb{P}(X \leq 1) &= \text{Pois}(1; 0.8) \overset{LOOKUP}{=} \boxed{0.80879} \\
\mathbb{P}(X \leq 2) &= \text{Pois}(2; 0.8) \overset{LOOKUP}{=} \boxed{0.95258} \\
\mathbb{P}(X = 2) &= \mathbb{P}(X \leq 2) - \mathbb{P}(X \leq 1) = 0.95258 - 0.80879 = \boxed{0.14379} \\
\mathbb{P}(X > 2) &= 1 - \mathbb{P}(X \leq 2) = 1 - 0.95258 = \boxed{0.04742}
\end{aligned}$$

Textbook Logistics for Section 3.6

- Difference(s) in Notation:

CONCEPT	TEXTBOOK NOTATION	SLIDES/OUTLINE NOTATION
Probability of Event	$P(A)$	$\mathbb{P}(A)$
Support of a r.v.	"All possible values of X"	Supp(X)
pmf of a r.v.	$p(x)$	$p_X(k)$
Expected Value of a r.v.	$E(X)$	$\mathbb{E}[X]$
Variance of a r.v.	$V(X)$	$\mathbb{V}[X]$
Poisson parameter	μ	λ
Poisson cdf	$F(x;\mu)$	Pois$(x;\lambda)$

Textbook Logistics for Section 3.6

- Skip ”The Poisson Distribution as a Limit” section (pg 132-133)
  - This is will be covered in Chapter 5 where it’s more appropriate.

- Ignore ”little-Oh” notion for assumptions of a Poisson Process (pg 134)
  - ”Little-Oh” notation involves limits that can be hard to interpret:

$$\text{e.g.} \quad (\Delta t)^3 = o(\Delta t) \text{ since } \lim_{\Delta t \to 0} \frac{(\Delta t)^3}{\Delta t} = 0$$

  - ”Little-Oh” notation is used for approximations, but it can be subtle.
  - Hence, ”Little-Oh” notation will never be considered in this course.
  - The way I framed the Poisson assumptions is ”loose” compared to using ”little-oh” notation, but for a first course in statistics it is perfectly sufficient.

Fin.