Review (2.5)

- Moment-generating function (m.g.f.) and its properties
- Calculating mean and variance through m.g.f.
- A note about discrete distribution

Today's Lecture(2.6)

- Poisson Processes, Poisson distribution.
- Mean, variance and m.g.f. of Poisson distribution.
- Poisson Approximation to Binomial distribution.

Poisson distribution

Examples:

1. Number of telephone calls arriving at a switchboard between 3pm and 5pm.
2. Number of defects in a 100-foot roll of aluminum screen that is 2 feet wide.
3. Number of road accidents in a year in US.

Poisson Process The Poisson process counts the number of events occurring in a fixed time or space, when

- the number of events occurring in non-overlapping intervals are independent.
- events occur at a constant average rate of λ per unit time.
- events cannot occur simultaneously.

Definition: The random variable X has a *Poisson distribution* if its p.m.f. is of the form

$$f(x) = \frac{\lambda^x}{x!} e^{-\lambda}, \quad \text{for } x = 0, 1, 2, \cdots$$

where $\lambda > 0$.

$\sum f(x) = 1$

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

It's a valid probability model.

Connections between Poisson process and Poisson distribution

Let X_t be the number of events to occur in time t (units). Then $X_t \sim$ Poisson(λt), and

$$P(X_t = x) = \frac{(\lambda t)^x}{x!} e^{-\lambda t} \quad \text{for } x = 0, 1, 2, \cdots,$$

The parameter λ is called the rate of Poisson Process.

Example: Telephone calls enter a college switchboard on the average of two every 3 minutes. Let X be the number of calls in a 9 minutes period. Assuming it is a Poisson process, what is the distribution of X?

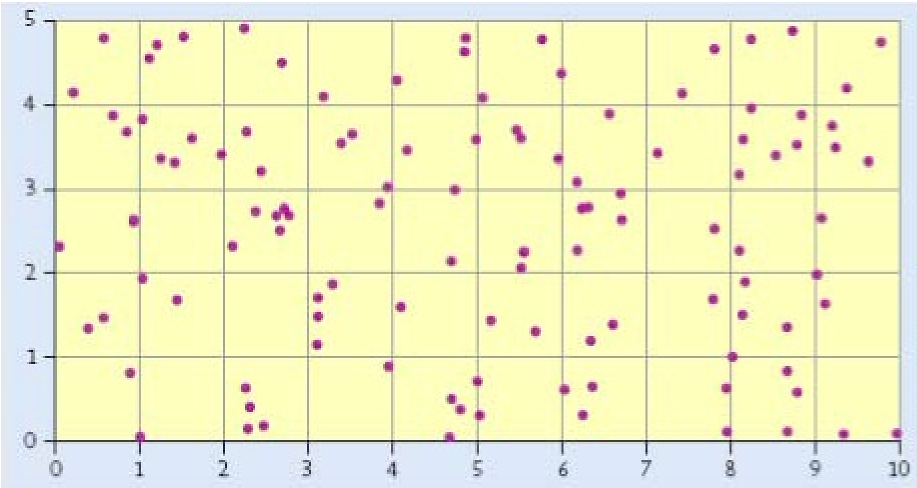
$\lambda = 2/3$

$t = 9$ min

$X \sim \text{Pois}(\lambda t) = \text{Pois}(6)$

Note: for a Poisson process in space, let $X_A = \#$ events in area of size A. Then $X_A \sim$ Poisson(λA).

Example: Dandelions seeds are wind-spread. If we divide a natural lawn into 1 ft^2 grids, we can count how many dandelions are in each grid. The number of dandelions in the area of size A is given by a Poisson distribution.



(i) independence of dandelions : the presence of one dandelion in a grid does not make the presence of another more or less likely.
(ii) homogeneity of grid: each grid is equally susceptible of containing dandelions.

Example: X_A = number of raisins in a volume A of current bun.

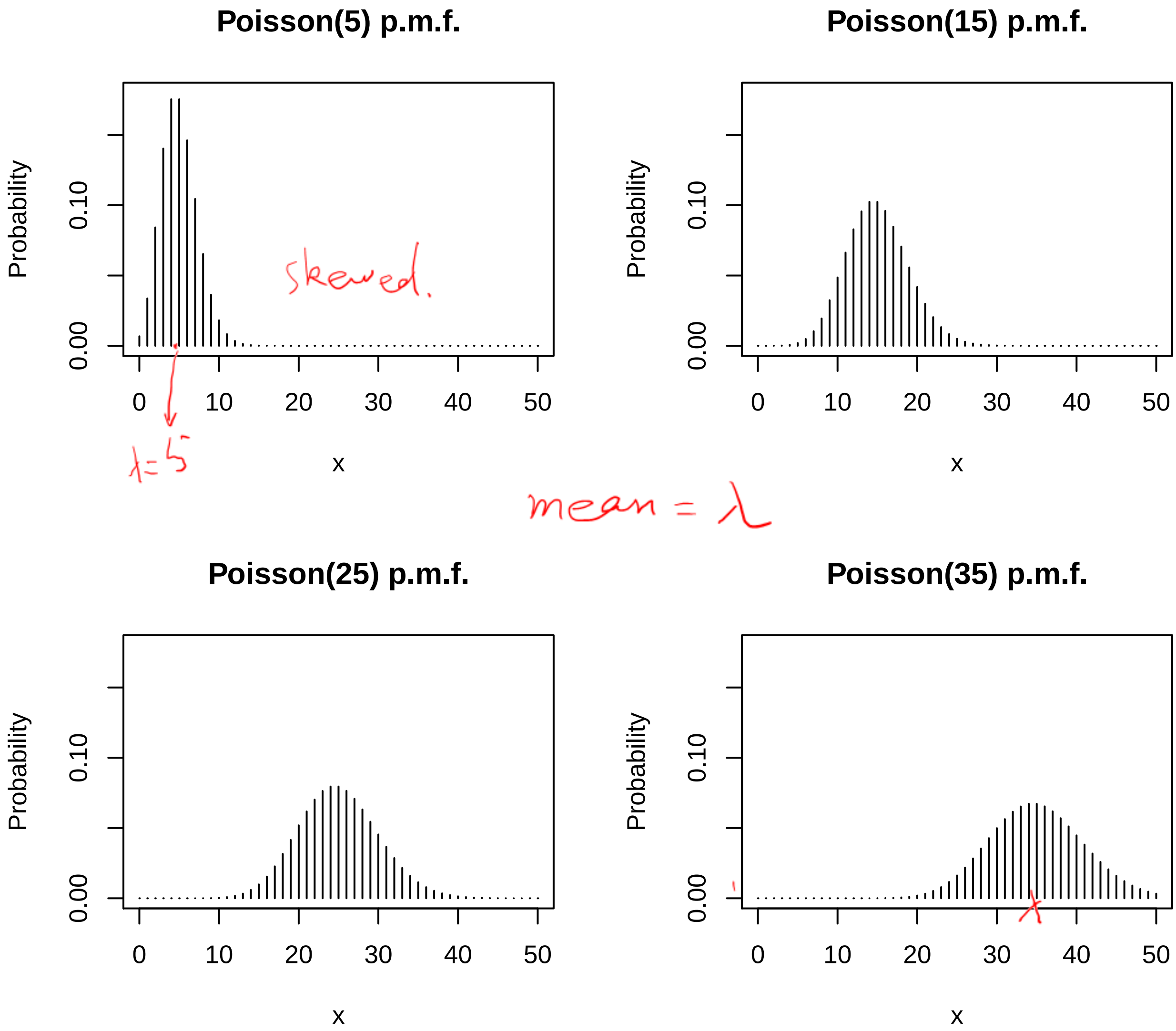


Figure 1: Effect of changing λ on Poisson distributions

Probability histogram of Poisson(λ)

Mean and Variance of Poisson distribution

$$f(x) = \frac{\lambda^x e^{-\lambda}}{x!}, \quad x = 0, 1, 2, 3, \cdots,$$

- Step 1: Find the m.g.f. $M(t)$

$M(t) = Ee^{tX}$

$= \sum_{x=0}^{\infty} e^{tx} \cdot e^{-\lambda} \cdot \frac{\lambda^x}{x!}$

$= \left[\sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{x!}\right] \cdot e^{-\lambda}$

$= e^{\lambda(e^t - 1)}$

- Step 2: Find the first and second derivatives
  - $M'(t) = \lambda e^t e^{\lambda(e^t-1)}$.
  - $M''(t) = (\lambda e^t)^2 e^{\lambda(e^t-1)} + \lambda e^t e^{\lambda(e^t-1)}$.
- Step 3: Find μ, σ^2:

$$\begin{aligned} \mu &= M'(0) = \lambda, \\ \sigma^2 &= M''(0) - [M'(0)]^2 = (\lambda^2 + \lambda) - \lambda^2 = \lambda \end{aligned}$$

Note: It makes sense for $E(X) = \lambda$, by definition, λ is the average number of events per unit time in the Poisson process.

Note: For Poisson (λ)

both mean and variance $= \lambda$.

Example

The number of deer crossing a road at night during mating season in a particular rural area can be modeled with a Poisson distribution. A local survey conducted over 4 nights found a total of 20 deer crossings. Based on this information, what is the probability that ... $\lambda = \frac{20}{4} = 5$

- fewer than three deer would cross on a given night during mating season in this area?

  X = # of deers will cross on a given night

  $X \sim Pois(5)$

  $P(X<3) = P(X=0,1,2) = \underbrace{e^{-5} \cdot \frac{5^0}{0!} + e^{-5} \cdot \frac{5^1}{1!} + e^{-5} \cdot \frac{5^2}{2!}}_{=0.125}$

  To compute this probability using the Poisson distribution, we need to know λ. In this case $\lambda = 20/4 = 5$ deer crossings per night

- no deer would cross on the given night.

  $P(X=0) = e^{-5} \cdot \frac{5^0}{0!} = e^{-5} \approx 0.0067$

- at least three deers would cross on the given night.

  $P(X \geq 3) = 1 - P(X<3) = 1 - 0.125$
  $= 0.875$

Example

Example: In the decade of the 1980s, the number X of cases of diphtheria (an acute febrile contagious disease) reported each year in the United States followed a Poisson distribution with mean 2.4. What is the probability that

$\lambda = 2.4$

- At most three were reported?

  $P(X \leq 3) = F(3) = 0.779$

  cdf $P(X \leq x) = F(x)$

- At least three were reported?

  $P(X \geq 3) = 1 - F(2)$

- Exactly three were reported?

  $P(X=3) = F(3) - F(2)$

x	0	1	2	3	4	5	6
$F(x)$	0.091	0.308	0.570	0.779	0.904	0.964	0.988
x	7	8	9	10	11	12	
$F(x)$	0.997	0.999	1.000	1.000	1.000	1.000	

Poisson Approximation to Binomial distribution

If the distribution of X is $b(n, p)$, then Poisson distribution with $\lambda = np$ can approximate it well if n is large and p is small.

Example: Two dice are rolled 100 times, and the number of double sixes, X, is counted. The distribution of X is binomial with $n = 100$ and $p = \frac{1}{36}$. Since n is large and p is small, we can approximate the binomial probabilities by Poisson probabilities with $\lambda = np = 2.78$.

$b(100, \frac{1}{36})$ $Pois(\frac{100}{36})$

x	Binomial Prob	Poisson Approximation
0	0.0596	0.0620
1	.1705	.1725
2	.2414	.2397
3	.2255	.2221
4	.1564	.1544
5	.0858	.0858
6	.0389	.0398
7	0.0149	0.0158
8	0.0050	0.0055
9	0.0015	0.0017
10	0.0004	0.0005
11	0.0001	0.0001

In general, the approximation is quite accurate if $n \geq 20$ and $p \leq 0.05$ or if $n > 100$ and $p \leq 0.10$.

In excel, use

=POISSON(x, λ, 0) to get P(X = x), p.m.f

=POISSON(x, λ, 1) to get P(X ≥ x), the cdf.

Example

The probability of being dealt a royal straight flush (ace, king, queen, jack, and ten of the same suit) in poker is about 1.54×10^{-6}. Suppose that an avid poker player sees 100 hands a week, 52 weeks a year, for 20 years. What is the probability that she never sees a royal straight flush dealt? What is the probability that she sees two royal straight flushes dealt? (Hint: use Poisson approximation)

$X \sim b(n, p)$

with $n = 100 \times 52 \times 20 = 104000$

$p = 1.54 \times 10^{-6}$

$P(X=0) = \binom{n}{0} p^0 (1-p)^n = (1 - 1.54 \times 10^{-6})^{104000}$ complicated

$P(X=2) = \ldots$

Use poisson approximation!

$\lambda = n \times p \doteq 104000 \times 1.54 \times 10^{-6} = 0.16$

$Y \sim Pois(0.16)$

$P(Y=0) = e^{-0.16} \times \frac{0.16^0}{0!} \doteq 0.85$

$P(Y=2) = e^{-0.16} \times \frac{0.16^2}{2!} \doteq 0.01$