DISCRETE RANDOM VARIABLES

The Poisson distribution

Poisson random variables

- Some random variables arise from counting – typically the occurrence of some phenomenon during a fixed amount of time.
- Examples:
  - Number of phone calls during one hour to a poison control hotline
  - Number of live births at a hospital in a week
  - Number of traffic accidents at an intersection in a month
- It doesn't have to be time (but it often is):
  - Number of weeds in a section of farmland
  - Number of lesions in a brain region

Derivation of the Poisson distribution

- We are counting the number of phone calls received in an hour.
- Let's break the hour up into many tiny intervals (think: intervals so small it's hard to conceive of there being more than one call per interval). (one second?)
- Suppose that for any given tiny interval: $P(one\ phone\ call) = p$.
- So in that tiny interval, $P(zero\ phone\ calls) = 1 - p$.
- (This then implies that for any tiny interval $P(more\ than\ one\ phone\ call) = 0$...)

- Suppose that phone calls coming are independent of each other.
- If Y is the total number of calls in an hour, and if there are n tiny intervals, then:

$$Y \sim Binom(n, p)$$

- Expected number of calls in the hour is np (right?).
- What if I cut each tiny interval in half? (Number of intervals is $n' = 2n$)
- Expected number of calls in the hour is still $np = n'p' = (2n)(\frac{p}{2})$.
- Let λ be the expected number of calls in the hour. (This should not depend on how tiny the intervals are.)
- So for any choice of $\boxed{n, p = \lambda/n.}$

- Since Y is binomial,
- $P(Y = y) = \binom{n}{y} p^y (1-p)^{n-y}$
- Let's think about the tiny intervals getting ever more tiny…

$$P(Y=y) = \binom{n}{y} p^y (1-p)^{n-y}$$

$$= \frac{n!}{y!\,(n-y)!}\, p^y (1-p)^{n-y}$$

$np = \lambda$

$p = \frac{\lambda}{n}$

FACT Appendix A.6

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

y terms

$$\lim_{n\to\infty} P(Y=y) = \lim_{n\to\infty} \frac{n(n-1)\ldots(n-y+1)}{y!}\left(\frac{\lambda}{n}\right)^y \left(1-\frac{\lambda}{n}\right)^{n-y}$$

y terms

$$\lim_{n\to\infty} \frac{n}{y}\,\frac{n-1}{y-1}\,\frac{n-2}{y-2}\cdots\frac{n-y+1}{1}\,\frac{\lambda^y}{n^y}\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-y}$$

y terms

$$\lim_{n\to\infty} \frac{n}{y\,n}\,\frac{n-1}{(y-1)\,n}\,\frac{n-2}{(y-2)\,n}\cdots\frac{n-y+1}{1\cdot n}\,\lambda^y \underbrace{\left(1-\frac{\lambda}{n}\right)^n}_{\to e^{-\lambda}\ \text{(FACT)}}\underbrace{\left(1-\frac{\lambda}{n}\right)^{-y}}_{\to 1}$$

$\frac{n}{y\,n} \to \frac{1}{y}$, $\frac{n-1}{(y-1)\,n} \to \frac{1}{y-1}$, $\frac{n-2}{(y-2)\,n} \to \frac{1}{y-2}$, $\frac{n-y+1}{1\cdot n} \to \frac{1}{1}$

$$\boxed{= \frac{\lambda^y}{y!}\, e^{-\lambda}}$$

Poisson distribution

- The Poisson distribution can be defined as the limit of a binomial distribution.
- A random variable Y has a POISSON DISTRIBUTION with rate parameter λ probabilities given by

$$P(Y = y) = \frac{\lambda^y}{y!}e^{-\lambda}, \qquad y = 0, 1, 2, \dots .$$

- Check: Is this a valid distribution function?

$$\sum_{y=0}^{\infty} P(Y = y) = \sum_{y=0}^{\infty} \frac{\lambda^y}{y!}e^{-\lambda} = e^{-\lambda}\sum_{y=0}^{\infty} \frac{\lambda^y}{y!} = e^{-\lambda}e^{\lambda} = 1$$

e^{λ}

(See Appendix A1.11)

Example 9

Poisson $\quad \lambda = 2.6$

- Accidents occur at a workplace at a rate of 2.6 per month. What is the probability that there will be 4 accidents next month?

counting

Model as $\quad Y \sim \text{Poisson}(2.6)$

$$P(Y=4) = \frac{(2.6)^4}{4!} e^{-2.6} = 0.113$$

Example 9 (continued)

- What is the probability there are between 3 and 6 accidents next month?

$P(3 \le Y \le 6) = P(Y=3) + P(Y=4) + P(Y=5) + P(Y=6)$

- See Appendix 3 Table 3 or use `ppois()` in `R`

$$P(3 \le Y \le 6) = P(Y \le 6) - P(Y \le 2)$$

$P(Y \le 6)$:

```
> ppois(q=6,lambda=2.4)
[1] 0.9884059
```

$P(Y \le 2)$:

```
> ppois(q=2,lambda=2.4)
[1] 0.5697087
```

$P(3 \le Y \le 6) = 0.988 - 0.5697 = 0.418$

Mean and variance of the Poisson distribution

- If Y is a Poisson random variable with rate parameter λ then

$$E[Y] = \lambda; \qquad Var(Y) = \lambda$$

$$E[Y] = \sum_{y=0}^{\infty} y \, P(Y=y)$$ def. of exp.

$y=0$ term $=0$

$$= \sum_{y=1}^{\infty} \cancel{y} \frac{\lambda^y}{\cancel{y!}_{(y-1)!}} e^{-\lambda} = \sum_{y=1}^{\infty} \frac{\lambda^y}{(y-1)!} e^{-\lambda}$$

change of variables

$x = y-1 \qquad y = x+1$

$y=1 \rightarrow x=0$

"$y=\infty$" $\rightarrow$ "$x=\infty$"

$$= \sum_{x=0}^{\infty} \frac{\lambda^{x+1}}{x!} e^{-\lambda} = \lambda \underbrace{\sum_{x=0}^{\infty} \frac{\lambda^x}{x!} e^{-\lambda}}_{=1}$$

$$= \lambda$$

Some notes about the Poisson distribution

- If $Y_1 \sim Poisson(\lambda_1)$ and $Y_2 \sim Poisson(\lambda_2)$ and Y_1 and Y_2 are independent, then
$$Y_1 + Y_2 \sim Poisson(\lambda_1 + \lambda_2)$$
- If λ is the rate of occurrences for one time period, then the distribution of the number of occurrences for m time periods is Poisson($m\lambda$).
- Note that the above requires that the rate be the same for all time periods.
- Just because you are counting some occurrences doesn't necessarily mean the random variable has a Poisson distribution!