Stat 401 - 11/3/17

Sec 3.1 - continued.

(5) Trinomial Dist. (special case of multinomial when $k=3$).

Let $X_1 = X$ $X_2 = Y$ $X_3 = n - X - Y$

Joint pmf of $X + Y$:

$$p(x,y) = \frac{n!}{x!\,y!\,(n-x-y)!} p_1^{x} p_2^{y} p_3^{n-x-y}$$

where $x+y \le n$, x,y nonneg integers.
$p_1 + p_2 + p_3 = 1$

Conditional pmf of Y given $X = x$ is

$$p_{2|1}(y|x) = \begin{cases} \frac{(n-x)!}{y!\,(n-x-y)!} \left(\frac{p_2}{1-p_1}\right)^{y} \left(\frac{p_3}{1-p_1}\right)^{n-x-y} & \text{for } y = 0, 1, \ldots, n-x \\ 0, & \text{o.w.} \end{cases}$$

Binom pmf. $\binom{n}{x} p^x (1-p)^{n-x}$

This implies that $Y | X = x \sim Bin\left(n-x, \frac{p_2}{1-p_1}\right)$

and $E[Y|x] = (n-x)\left(\frac{p_2}{1-p_1}\right)$

$E[X|y] = (n-y)\left(\frac{p_1}{1-p_2}\right)$.

We can also see this relationship through the correlation coefficient

$$\rho = -\sqrt{\frac{p_1 p_2}{(1-p_1)(1-p_2)}} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

⑥ Geometric Distribution.

Let Y = # of failures before the 1st success.

ex)

S	← 0 failures	$y=0$
FS	← 1 failure	$y=1$
FFS	← 2 failures	$y=2$

PMF

$$P_Y(y) = \begin{cases} p(1-p)^y, & y=0,1,2,\dots \\ 0, & \text{o.w.} \end{cases}$$

Assume independent trials.

$\boxed{Q}$ What if we want exactly r successes? (stop when obtain r^{th} success).

No fixed n.

7 Negative Binomial Dist.

Y = # of failures before the r^{th} success

(# trials required to produce exactly r successes $= Y + r$)

r = fixed positive integer.

ex) $r = 2$. (want 2 successes total
last trial = 2nd success)

SS ← 0 failures $y=0$
SFS ← 1 failure $y=1$

FSS $y=1$

SFFFFFFFS $y=7$

$$\underbrace{\binom{y+r-1}{r-1}}_{\text{ways to arrange failures}} p^{r-1}(1-p)^{y} \quad p$$

pmf of Y is

$$P_Y(y) = \begin{cases} \binom{y+r-1}{r-1} p^r (1-p)^y & \text{for } y = 0,1,2,\ldots \\ 0 & \text{, o.w.} \end{cases}$$

MGF

$$M_Y(t) = p^r \left[1-(1-p)e^t\right]^{-r}$$

for $t < -\log(1-p)$

ln

In Stat 381:

Neg Binom $X = \#$ trials until k^{th} success.

$$P_X(x) = \begin{cases} \binom{x-1}{k-1} p^k (1-p)^{x-k} & x = k, k+1, \ldots \\ 0 & \text{, o.w.} \end{cases}$$

Geom $X = \#$ trials until 1^{st} success

$$P_X(x) = \begin{cases} (1-p)^{x-1} p & , x = 1,2,\ldots \\ 0 & , \text{o.w.} \end{cases}$$

Sec 3.2 - Poisson Dist.

Recall a special series.

$$1 + m + \frac{m^2}{2!} + \frac{m^3}{3!} + \dots = \sum_{x=0}^{\infty} \frac{m^x}{x!}$$

→ converges for all m

→ converges to e^m

def) Poisson Dist.

$$p_X(x) = \begin{cases} \dfrac{\mu^x e^{-\mu}}{x!}, & x = 0, 1, 2, \dots \\ 0, & \text{o.w.} \end{cases}$$

where $\mu > 0$

μ = parameter → it sets the distribution → determines what it looks like.

$X \sim \text{Pois}(\mu)$

Models # of events happening in a fixed period of time.

$$M(t) = e^{\mu(e^t - 1)} \quad \text{for } -\infty < t < \infty$$

$$\mathbb{E}[X] = \mu = \text{Var}[X].$$

Purpose of μ: this is a parameter that describes the intensity.

<u>Probability Calcs.</u>

$$\mathbb{P}[X = k] = p_X(k)$$

$$\mathbb{P}[X \leq k] = F_X(k)$$

$$= \mathbb{P}[X = k] + \mathbb{P}[X = k-1] + \dots + \mathbb{P}[X = 0].$$

$$\mathbb{P}[X \geq k] = 1 - \mathbb{P}[X < k]$$

$$= 1 - \mathbb{P}[X \leq k-1]$$

<u>ex)</u> $X \sim \text{Pois}(2)$ μ

$$\text{PMF: } p(x) = \begin{cases} \dfrac{2^x e^{-2}}{x!}, & x = 0, 1, 2, \dots \\ 0, & \text{o.w.} \end{cases}$$

$$\mathbb{E}[X] = \mathrm{Var}[X] = 2$$
$$\mu = \sigma^2 = 2$$

$$\begin{aligned}
\mathbb{P}[1 \le X] &= \mathbb{P}[X \ge 1] \\
&= 1 - \mathbb{P}[X < 1] \\
&= 1 - \mathbb{P}[X \le 0] \\
&= 1 - \mathbb{P}[X = 0] \\
&= 1 - p_X(0) \\
&= 1 - \frac{2^0 e^{-2}}{0!} \\
&= 1 - e^{-2} \\
&= 1 - 0.135
\end{aligned}$$

ex) $\mathbb{P}[X \le 4] = 0.947$

$$\begin{aligned}
\mathbb{P}[X = 4] &= \mathbb{P}[X \le 4] - \mathbb{P}[X \le 3] \\
&= 0.947 - 0.857
\end{aligned}$$

ex) $M_X(t) = e^{4(e^t - 1)}$

$\mu = 4$

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

$X \sim Pois(4)$

Poisson Process

Setup:

→ Let $g(x, w)$ denote the prob. of x changes in each interval of length w.

ex) 3 claims in 2 days
$g(3, 2)$

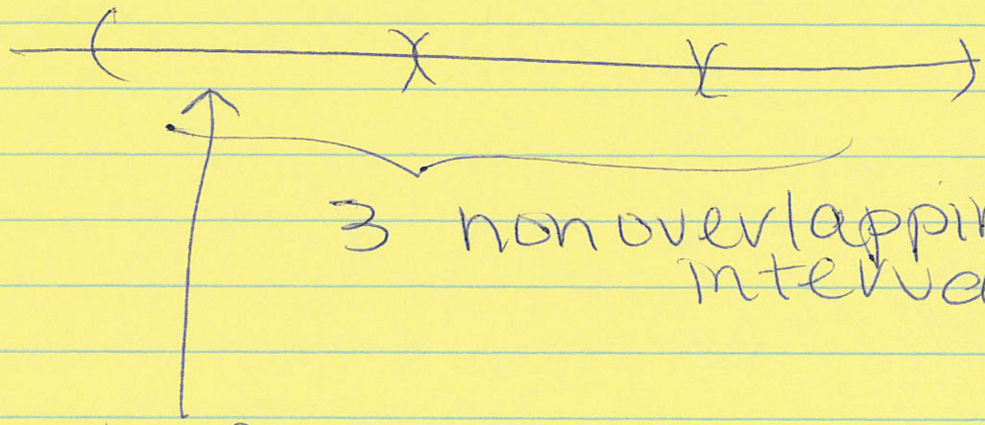
define $g(0, 0) = 1$

Poisson Postulates

① the prob. of 1 change in an interval h is equal to the length of interval + a little bit.

2. the prob. of 2 or more changes in an interval h is negligible.

3. The numbers of changes in non-overlapping intervals are independent.



3 nonoverlapping intervals

\# of changes in interval 1 does not impact # of changes in any other interval.

In summary:

Take an interval of length w. Let the # of changes in an interval of length 1 be denoted by λ.

then the # of changes in
ω (length of time interested in)
is given by

$$\lambda\omega$$

corresponding parameter
for Poisson dist.

<u>ex)</u> 5 accidents / hr.

$\lambda = 5$ $\quad \omega = 1$

Prob there are exactly 4 accidents
in 1 hr.

$\mu = 5$ $\quad \mathbb{P}[X=4]$

$X \sim Pois(5)$.

Prob there are exactly 4 accidents
in 2 hours.

$\lambda = 5$ $\quad \omega = 2$

$\lambda\omega = 10$ $\quad Y \sim Pois(10)$

$\mathbb{P}[Y = 4]$

<u>Thm</u> Suppose $X_1, \ldots, X_n$ are indep RVs.

Take $X_i \sim Pois(\mu_i)$

then $Y = \sum_{i=1}^{n} X_i \sim Pois\left(\sum_{i=1}^{n} \mu_i\right)$