Lecture 15

Longhai Li, October 28, 2021

plan:

✓ 1. Bernoulli, Binomial, Multinomial

Geometric, NB, poisson Sec 3.1, 3.2

2. Gamma, χ^2, exponential.

β dis. sec 3.3

3. Normal distribution 1. [illegible]

$\bar{X} \perp S^2$. $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$

$(n-1)S^2/\sigma^2 \sim \chi^2$

Bernoulli (Bern (p)) $p \in [0,1]$

$$X \sim \begin{array}{c|cc} x & 0 & 1 \\ \hline p & q & p \end{array}, \text{ where } q = 1 - p$$

Example:

$Y_i \sim \text{Bern}(p_i)$

$$p_i = \frac{e^{\boxed{x_i}\beta}}{1 + e^{x_i \beta}}$$ logistic regression

$X \sim \mathrm{Bern}(p)$

is distributed as

$$E(X) = 0 \cdot q + 1 \cdot p = p$$

$$E(X^2) = 0^2 q + 1^2 \cdot p = p$$

$$V(X) = \mathbf{E}(X^2) - [E(X)]^2 = p - p^2$$

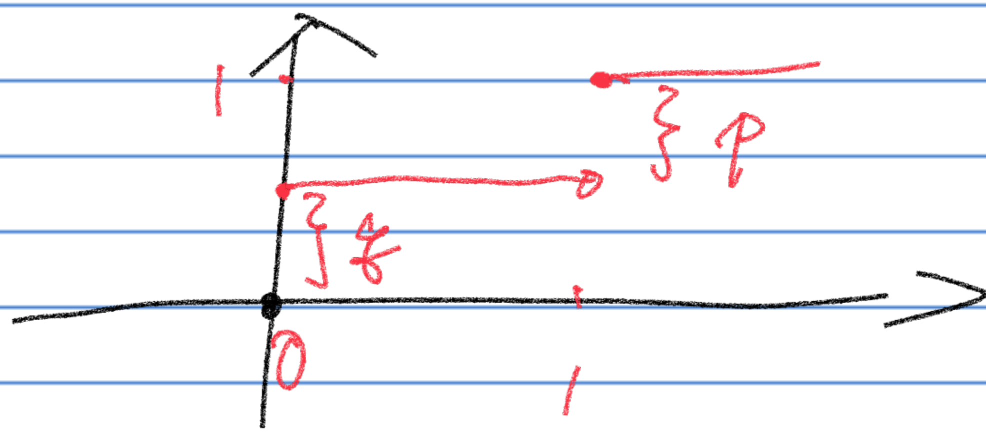
$$= p \cdot q$$

$$M_X(t) = E\left(e^{tX}\right)$$

$$= \underbrace{e^{t\cdot 0}}_{1} \cdot q + e^{t\cdot 1} \cdot p$$

$$= q + p \cdot e^{t}$$

C.D.F.



Binomial distribution

n independent trials. each succeeds with prob. p.

$X_1, X_2, \cdots, X_n$ $\quad X_i = \begin{cases} 1, & \text{success} \\ 0, & \text{failure} \end{cases}$

$X_i \sim \text{Bern}(p)$

$X_1, \cdots, X_n \overset{\text{IID}}{\sim} \text{Bern}(p)$

Let $Y = \sum_{i=1}^{n} X_i \sim \text{Binomial}(n, p)$

p.m.f.

$$p_Y(y) = \binom{n}{y} \cdot p^y q^{n-y}, \text{ for } y = 0, 1, \cdots, n$$

$P(Y = y)$

$\boxed{1} \cdots \boxed{0} \cdots \boxed{1}$

y positions are 1

$$Y = 2 \left\{ \begin{array}{l} 1\,0\,0\,0\,0\,0\,1 \\ 1\,1\,0\,0\,0\,0\,0 \\ \vdots \\ 0\,0\,0\,0\,0\,1\,1 \end{array} \right\} \binom{n}{2}$$

Binomial Formula

Combination

$$(a+b)^n = \sum_{y=0}^{n} \binom{n}{y} a^y b^{n-y}$$

$$1 = (p+q)^n = \sum_{y=0}^{n} \binom{n}{y} p^y \cdot q^{n-y}$$

$$\sum_{y=0}^{n} P_Y(y) = 1$$

$$E(Y) = E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i)$$

$$= \sum_{i=1}^{n} \cdot p = np$$

$$V(Y) = V\left(\sum_{i=1}^{n} X_i\right)$$

$$= \sum_{i=1}^{n} V(X_i) = \sum_{i=1}^{n} (p \cdot q)$$

$$= n \cdot p \cdot q$$

M.G.F.

$$M_Y(t) = E\left(e^{t\sum_{i=1}^{n} X_i}\right)$$

$$= \left[M_{X_i}(t)\right]^n$$

$$= \left[q + pe^t\right]^n$$

Multinomial distribution.

n indep. trials, each having $k+1$ outcomes denoted by $1, 2, \dots, k, 0$, with success rates $p_1, p_2, \dots, p_k, p_0$, respectively.

Let $(X_1^{(i)}, X_2^{(i)}, \dots, X_k^{(i)})$ be the outcome of the ith trial

$$X_1^{(i)} = \begin{cases} 1, & \text{if } i\text{th outcome is } 1 \\ 0, & \text{o.w.} \end{cases}$$

$$X_2^{(i)} = \begin{cases} 1, & \text{if } i\text{th outcome is } 2 \\ 0, & \text{o.w.} \end{cases}$$

$$\vdots$$

$$X_k^{(i)} = \begin{cases} 1, & \text{if } i\text{th outcome is } k \\ 0, & \text{o.w.} \end{cases}$$

$$Y_1 = \sum_{i=1}^{n} X_1^{(i)}, \quad Y_2 = \sum_{i=1}^{n} X_2^{(i)}$$

$$\ldots \quad Y_k = \sum_{i=1}^{n} X_k^{(i)}$$

Y_j : # of trials with outcome j

$$(Y_1, \ldots, Y_k) \sim \text{Multinomial}(n, p_1 \ldots p_k)$$

Joint P.M.F.

$P(y_1, \ldots, y_k)$

$$= \frac{n!}{y_1! \cdots\cdots y_k!\, y_0!} \cdot p_1^{y_1}\, p_2^{y_2} \cdots p_k^{y_k}\, p_0^{y_0}$$

$$\underbrace{\boxed{1} \cdots \boxed{1}}_{y_1},\ \underbrace{\boxed{2} \cdots \boxed{2}}_{y_2},\ \cdots\cdots,\ \underbrace{\boxed{k} \cdots \boxed{k}}_{y_k}$$

$$\underbrace{\boxed{0} \cdots\cdots \boxed{0}}_{n-(y_1+\cdots+y_k) = y_0} \qquad \binom{n}{y_1\ y_2 \cdots y_k,\ y_0}$$

where $y_0 = n - \sum_{i=1}^{k} y_i$, $p_0 = 1 - \sum_{i=1}^{k} p_i$

$\sum_{i=1}^{k} y_i \le n$.

M.G.F.

$$M_{Y_1, \dots, Y_k}(t_1, \dots, t_k)$$

$$= E\left(e^{t_1 Y_1 + t_2 Y_2 + \dots + t_k Y_k}\right)$$

$$= E\left(e^{\sum_{i=1}^{n} [t_1 X_1^{(i)} + \dots + t_k X_k^{(i)}]}\right)$$

$$= \prod_{i=1}^{n} E\left(e^{t_1 X_1^{(i)} + \dots t_k X_k^{(i)}}\right)$$

$$= \left(p_0 + p_1 e^{t_1} + p_2 e^{t_2} + \cdots + p_k e^{t_k} \right)^n$$

e^{tx}	$(X_1^{(i)}, \ldots, X_k^{(i)})$	prob
p_0	$(0, 0, \ldots, 0)$	p_0
e^{t_1}	$(1, 0, \ldots, 0)$	p_1
e^{t_2}	$(0, 1, \ldots, 0)$	p_2
$\vdots$	$\vdots$	
e^{t_k}	$(0, 0, \ldots, 1)$	p_k

$$M_{Y_1,\dots,Y_k}(t_1,\dots,t_k)$$

$$= E\left(e^{t_1Y_1+t_2Y_2+\dots+t_kY_k}\right)$$

$$M_{Y_1,Y_2}(t_1,t_2)$$

$$= M_{Y_1,\dots,Y_k}(t_1,t_2,0,\dots,0)$$

$$M_{Y_i}(t_i) = M_{Y_1,\dots,Y_k}(0,0,0,\dots,t_i,0,\dots,0)$$

$$M_{Y_1,\dots,Y_k}(t_1,\dots,t_k)$$

$$= \left(p_0 + p_1 e^{t_1} + \dots + p_k e^{t_k}\right)^n$$

$$M_{Y_i}(t_i) = \Big(\underbrace{p_0 + \sum_{j\neq i} p_j}_{q} + \underbrace{p_i}_{p} e^{t_i}\Big)^n$$

$$Y_i \sim \text{Binomial}(n, p_i) \quad \checkmark$$

$$M_{Y_i, Y_j}(t_i, t_j)$$

$$= \Big(\underset{\substack{\uparrow \\ p_0}}{\sum_{\substack{l \neq i, \\ l \neq j}} p_l} + \underset{\substack{\uparrow \\ p_i}}{p_i} e^{t_i} + \underset{\substack{\uparrow \\ p_j}}{p_j} e^{t_j} \Big)^n$$

$$(Y_i, Y_j) \sim \text{Multinomial}(n, \; p_i, \; p_j)$$

Conditional distribution.

$$Y_2, Y_3, \ldots, Y_k \mid Y_1 = y_1 \sim ?$$

$$p(y_2, y_3, \ldots, y_k \mid Y_1 = y_1)$$

$$= \frac{p(y_1, y_2, \ldots, y_k)}{P_{Y_1}(y_1)}$$

$$= \frac{\dfrac{\cancel{n!}}{\cancel{y_1!}\cdots y_k!\, y_0!}\, \cancel{p_1^{y_1}}\, p_2^{y_2} \cdots p_k^{y_k}\, p_0^{y_0}}{\dfrac{\cancel{n!}}{\cancel{y_1!}\,(n-y_1)!}\, \cancel{p_1^{y_1}}\, (1-p_1)^{n-y_1}}$$

$$= \frac{(n-y_1)!}{y_2!\, y_3! \cdots y_k!\, y_0!} \left(\frac{p_2}{1-p_1}\right)^{y_2} \cdots \left(\frac{p_0}{1-p_1}\right)^{y_0}$$

Note: $n - y_1 = y_0 + y_2 + \cdots + y_k$

$$Y_2, Y_3, \ldots, Y_k \mid Y_1 = y_1$$

$$\sim \text{Multinomial}\left(n - y_1, \frac{p_2}{1-p_1}, \frac{p_3}{1-p_1}, \ldots, \frac{p_k}{1-p_1}\right)$$