Jan 30th

EX2 Poisson distribution

$P[X_i = x_i] = \frac{\lambda^{x_i}}{x_i!} e^{-\lambda}$

Likelihood function $\prod_{i=1}^{n} \frac{\lambda^{x_i}}{x_i!} e^{-\lambda}$

$$l(\lambda) = \sum_{i=1}^{n} \log\left(\frac{\lambda^{x_i}}{x!} e^{-\lambda}\right) = \sum_{i=1}^{n} (x_i \log\lambda - \lambda - \log x_i!) = \log\lambda \sum_{i=1}^{n} x_i - n\lambda - \sum_{i=1}^{n} \log(x_i!)$$

$$l'(\lambda) = \frac{1}{\lambda}\sum_{i=1}^{n} x_i - n = 0 \Rightarrow \lambda = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}$$

Ex3 Normal Case

$$f(x|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}$$

$$\log f(x|\mu,\sigma^2) = -\log\sigma - \log\sqrt{2\pi} - \frac{1}{2}(x-\mu)^2/\sigma^2$$

Then log likelihood $= \sum_{i=1}^{n}[-\log\sigma - \log\sqrt{2\pi} - \frac{1}{2}(x-\mu)^2/\sigma^2] = -n\log\sigma - n\log\sqrt{2\pi} - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$

$$\frac{l(\mu,\sigma^2)}{\partial\mu} + \frac{1}{2\sigma^2}\cdot 2\sum_{i=1}^{n}(x_i-\mu) = 0 \Rightarrow \sum_{i=1}^{n}(x_i-\mu) = 0 \Rightarrow \mu = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}$$

$$\frac{l(\mu,\sigma^2)}{\partial\sigma^2} = -\frac{n}{2}\frac{1}{\sigma^2} + \left(\frac{1}{\sigma^2}\right)^2\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2 = 0 \Rightarrow -\frac{n}{2}\sigma^2 + \frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2 = 0 \Rightarrow \sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$$

Ex4 Uniform distributition

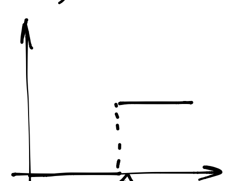
$x_i \sim Unif(a,b)$ x_i's are independent

not differentiable

$$f(x|a,b) = \frac{1}{b-a} I\{x\le b\} I\{x\ge a\} \Rightarrow \text{Likelihood} = \left(\frac{1}{b-a}\right)\prod_{i=1}^{n}\{x\le b\}\prod_{i=1}^{n}\{x\ge a\}$$

$$I\{x\le b\} = \begin{cases} 1 & \text{if } x\le b \\ 0 & \text{if } x>b \end{cases}$$

$$I\{x\ge a\} = \begin{cases} 1 & \text{if } x\ge a \\ 0 & \text{if } x<a \end{cases}$$



$\hat{b}_{MLE} = \max\{x_i\}$

$\hat{a}_{MLE} = \min\{x_i\}$

$\hat{a}_{MM} = \hat{m}_1 - \sqrt{3}\sqrt{\hat{m}_2^2 - \hat{m}_1^2}$

$\hat{b}_{MM} = \hat{m}_1 + \sqrt{3}\sqrt{\hat{m}_2^2 - \hat{m}_1^2}$

Estimating the cell prb of a multinomial distribution

Sps we have n types, For each type prob. for a give individual of the population falling into type i is p_i

$p_1 + p_2 + \cdots + p_n = 1$

Suppose I collect $x_1, x_2, \cdots, x_n$ sample

$$f(x_1, x_2, \cdots x_n | p_1, p_2, \cdots, p_m) = \frac{n!}{\prod_{i=1}^{m} x_i!}\prod_{i=1}^{m} p_i^{x_i}$$

Note $\sum_{i=1}^{m} P_i = 1$

Should use Lagrange multiplier

Let the sample size be n, then $x_1+x_2+\cdots+x_m=n$

$$G(\lambda, p_1, \cdots, p_m) = \log\left(\frac{n!}{\prod x_i!}\prod_{i=1}^{m} p_i^{x_i}\right) - \lambda\left(1-\sum_{i=1}^{m} p_i\right)$$

$$= \sum_{i=1}^{m} x_i \log p_i - \log\frac{n!}{\prod_{i=1}^{m} x_i!} - \lambda\left(1-\sum_{i=1}^{m} p_i\right)$$

$$\frac{\partial G}{\partial p_i} = x \cdot \frac{1}{p_i} + \lambda = 0 \Rightarrow p_i = -\frac{x_i}{\lambda}$$

$$\frac{\partial G}{\partial \lambda} = 0 \Rightarrow 1 - \sum_{i=1}^{m} p_i = 0 \Rightarrow \sum p_i = 1 \Rightarrow \sum -\frac{x_i}{\lambda} = 1 \Rightarrow \lambda = -\sum x_i$$
$$\Rightarrow \lambda = -n$$

Hardy–Weinberg Equilibrium

2 types of genes A a — $A \Downarrow$ Dominant, $a \Downarrow$ Ressessive

$p(A) = 1-\theta$
$p(a) = \theta$

AA	Aa	aa
$(1-\theta)^2$	$2\theta(1-\theta)$	θ^2

If I collect n persons & I call their geno

types $X_1 = \#$ of AA
$X_2 = \cdots$ Aa
$X_3 = \cdots$ aa

$x_1+x_2+x_3=n$

$$\text{likelihood} = \frac{n!}{\prod_{i=1}^{3} x_i!}\prod_{i=1}^{3} p_i^{x_i} = \frac{n!}{\prod_{i=1}^{3} x_i!}\left[(1-\theta)^2\right]^{x_1}\left[2\theta(1-\theta)\right]^{x_2}(\theta^2)^{x_3}$$

$$l(\theta) = \log\frac{n!}{\prod x_i!} + 2x_1\log(1-\theta) + x_2\log(2\theta(1-\theta)) + 2x_3\log\theta$$

$$l(\theta) = \log\frac{n!}{\prod x_i!} + (2x_1+x_2)\log(1-\theta) + (2x_3+x_2)\log\theta + x_2\log 2$$

$$l'(\theta) = -\frac{2x_1+x_2}{1-\theta} + \frac{2x_3+x_2}{\theta} = 0$$

Solve θ $\quad \theta(2x_1+x_2) = (1-\theta)(2x_3+x_2) \Rightarrow 2\theta(x_1+x_2+x_3) = 2x_3+x_2 \Rightarrow \theta = \frac{2x_3+x_2}{2n}$

Def: Consistency

If $\hat\theta$ is an estimcte of θ. We con $\hat\theta$ a cosistent estimator if $P[|\hat\theta-\theta|>\varepsilon] \to 0$ for any poisitive real number ε as $n\to\infty$

$\Updownarrow$ $\hat\theta \xrightarrow{P} \theta$ as $n\to\infty$

Thm A: Under sufficient smoothness conditions of $f(x|\theta)$ Then $\hat\theta_{MLE}$ is a constant estimator of θ.

Sketch of proof: Note that $\hat\theta_{MLE}$ is a zero point of $l'(\theta)$

Now $E\, l'(\theta) = E\left[\sum_{i=1}^{n} \log(f(x_i|\theta_0))\right]' = \left[E\sum_{i=1}^{n}(\log f(x_i|\theta_0))'\right] = E\left[\sum_{i=1}^{n} \frac{f'(x_i|\theta_0)}{f(x_i|\theta_0)}\right]$

$$= \sum_{i=1}^{n} E\left[\frac{f'(x_i|\theta_0)}{f(x_i|\theta_0)}\right] = \sum_{i=1}^{n}\int_{-\infty}^{\infty} \frac{f'(x|\theta_0)}{f(x|\theta_0)} \cdot f(x|\theta_0)\,dx = n\int_{-\infty}^{\infty} f'(x|\theta_0)dx$$

$$= n\left[\int_{-\infty}^{\infty} f(x|\theta_0)dx\right]' = n[1]' = 0 \qquad (1)$$

On the other hand:

$$Var\left[\frac{1}{n} l'(\theta_0)\right] = \frac{1}{n^2} Var(l'\theta_0) = \frac{1}{n^2} Var \sum_{i=1}^{n}[\log f(x_i|\theta_0)]' = \frac{1}{n} Var[\log - f(x_i|\theta_0)] \to 0 \; (2)$$

(1) & (2) implies $\frac{1}{n} l'(\theta_0) \simeq 0$ since $\frac{1}{n} l'(\hat{\theta}_{MLE}) = 0 \Rightarrow \hat{\theta}_{MLE} \xrightarrow{P} \theta_0$

Lemma $I(\theta) = E\left[\frac{\partial}{\partial\theta}\log f(X|\theta)\right]^2 \leftarrow$ fisher information

Then assume that $X \sim f$ density and $f(X|\theta)$ is smooth

Then $I(\theta) = -E\left[\frac{\partial}{\partial\theta^2}\log f(x|\theta)\right]$

First observe that

$$\int_{-\infty}^{\infty} \frac{\partial \log f(x|\theta)}{\partial\theta} f(x|\theta)\,dx = 0 \qquad (*)$$

Proof: $(*) = \int_{-\infty}^{\infty} \frac{f'(x|\theta)}{f(x|\theta)} f(x|\theta)\,dx$

$$= \int_{-\infty}^{\infty} f'(x|\theta)dx = \left[\int_{-\infty}^{\infty} f(x|\theta)dx\right]' = 1' = 0$$

take derivative w.r.t. (with respect to) θ on the LHS of $(*)$

$$\frac{\partial}{\partial\theta}[LHS] = \underbrace{\int_{-\infty}^{\infty} \frac{\partial}{\partial\theta^2}\log f(x|\theta) f(x|\theta)\,dx}_{B} + \underbrace{\int_{-\infty}^{\infty}\left[\frac{\partial}{\partial\theta}\log f(x|\theta)\right]\frac{\partial}{\partial\theta}f(x|\theta)dx}_{A}$$

$\boxed{[fg]' = f'g + fg'}$

Clearly $B = E\left[\frac{\partial}{\partial\theta^2}\log f(x|\theta)\right] = 0$

Q.E.D

Note that $\frac{\partial}{\partial\theta}\log f(x|\theta) = \frac{f'(x|\theta)}{f(x|\theta)}$

Therefore $f'(x|\theta) = f(x|\theta)\frac{\partial}{\partial\theta}\log f(x|\theta)$

$$\Rightarrow A = \int_{-\infty}^{\infty} \frac{\partial}{\partial\theta}\log f(x|\theta)\, f(x|\theta)\frac{\partial}{\partial\theta}\log f(x|\theta)\,dx$$

$$= \int_{-\infty}^{\infty}\left[\frac{\partial}{\partial\theta}\log f(x|\theta)\right]^2 f(x|\theta)\,dx = E\left[\frac{\partial}{\partial\theta}\log f(x|\theta)\right]^2 = I(\theta)$$

Thm B: If $f(x|\theta)$ is sufficiently smooth and the data X_i's are i.i.d.
Then $\sqrt{nI(\theta_0)}\,[\hat{\theta}_{MLE} - \theta_0] \longrightarrow N(0,1)$

Corollary: A $(1-\alpha)$% CI for θ_0 can be constructed as $\hat{\theta}_{MLE} \pm Z(1-\frac{\alpha}{2})\sqrt{nI(\hat{\theta}_{MLE})}$

cutoff for standard normal

Proof: Taylor expansion of $l'(\theta)$ centered at θ_0 we obtain

$$l'(\theta) = l'(\theta_0) + l''(\theta_0)(\theta - \theta_0) + \text{small term}$$

Let $\theta = \hat{\theta}_{MLE}$ we have

$$0 = l'(\hat{\theta}_{MLE}) = l'(\theta_0) + l''(\theta_0)(\hat{\theta}_{MLE} - \theta_0) + \text{small term} \Rightarrow -l'(\theta_0) \approx l''(\theta_0)(\hat{\theta}_{MLE} - \theta_0)$$

$$\sqrt{n}(\hat{\theta}_{MLE} - \theta_0) \approx -\frac{\frac{1}{\sqrt{n}}l'(\theta_0)}{\frac{1}{n}l''(\theta_0)}$$

Notice that $\frac{1}{n}l''(\theta_0) = \frac{1}{n}\sum_{i=1}^{n}\left[\frac{\partial \log f(x_i|\theta_0)}{\partial \theta^2}\right]$ According to the law of large numbers

$$\xrightarrow{P} E\left[\frac{\partial \log f(x|\theta_0)}{\partial \theta_0^2}\right] = -I(\theta_0)$$

first, note that $E[l'(\theta_0)] = 0$ (Check notes)

$$\frac{1}{\sqrt{n}}l'(\theta_0) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\partial \log f(x_i|\theta)}{\partial \theta}$$

Since $E[l'(\theta_0)] = 0$ Now we apply the <u>classic central limit theorem</u>

$$\frac{1}{\sqrt{n}}l'(\theta_0) \xrightarrow{P} N(0, \sigma^2) \text{ where } \sigma^2 = Var\left(\frac{\partial \log f(x|\theta_0)}{\partial \theta_0}\right) = I(\theta)$$

Hence $\sqrt{n}(\hat{\theta}_{MLE} - \theta_0) \approx \frac{N(0, I(\theta_0))}{I(\theta_0)} = N\left(0, \frac{1}{I(\theta_0)}\right)$

$$\Rightarrow \sqrt{nI(\theta_0)}\,(\hat{\theta}_{MLE} - \theta_0) \longrightarrow N(0,1)$$

Quiz 2 coverage:
p.223—226
p.255—274