9.2 Relative Efficiency

Ch. 8: we looked at estimating a parameter, θ, using a fct. of sample data, $\hat{\theta}$.

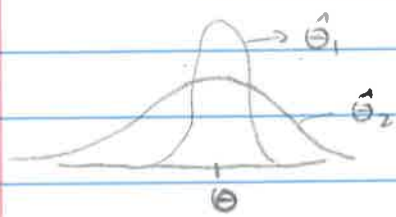
If $E(\hat{\theta}) = \theta$, we said $\hat{\theta}$ was unbiased for θ. Spse we have two unbiased estimators, $\hat{\theta}_1$ and $\hat{\theta}_2$. Does it matter which one we use?

Yes.

Def 9.1 Given two unbiased estimators $\hat{\theta}_1$ and $\hat{\theta}_2$ of a parameter θ, with variances $V(\hat{\theta}_1)$ and $V(\hat{\theta}_2)$, respectively, then the efficiency of $\hat{\theta}_1$ relative to $\hat{\theta}_2$, denoted $\text{eff}(\hat{\theta}_1, \hat{\theta}_2)$, is defined to be the ratio

$$\boxed{\text{eff}(\hat{\theta}_1, \hat{\theta}_2) = \frac{V(\hat{\theta}_2)}{V(\hat{\theta}_1)}}$$

$\hat{\theta}_1$ is better than $\hat{\theta}_2$ if $\text{eff}(\hat{\theta}_1, \hat{\theta}_2) > 1$ i.e. $V(\hat{\theta}_2) > V(\hat{\theta}_1)$



Example 9.1

let $y_1, \dots, y_n$ be a r.s. from $Y \sim \text{unif}(0, \theta)$

$\hat{\theta}_1 = 2\bar{Y}$ $\qquad \hat{\theta}_2 = \left(\frac{n+1}{n}\right) Y_{(n)}$

Find $\text{eff}(\hat{\theta}_1, \hat{\theta}_2)$

First, we need to determine if $\hat{\theta}_1$ and $\hat{\theta}_2$ are unbiased

$$E(Y) = \frac{0+\theta}{2} = \frac{\theta}{2}$$

$$E(2\cdot\bar{Y}) = 2E(\bar{Y}) = 2\cdot\theta/_2 = \boxed{\theta}$$ (i.e. $\hat{\theta}_1$ is unbiased)

$$V(2\bar{Y}) = 4V(\bar{Y}) = 4\cdot\frac{\sigma_Y^2}{n} = 4\,\frac{\frac{(\theta-0)^2}{12}}{n} = \boxed{\frac{\theta^2}{3n}}$$

$\hat{\theta}_2 = \frac{n+1}{n} Y_{(n)}$. We need the density of $Y_{(n)}$

$$F_{Y_{(n)}}(y) = P[Y_{(n)} \le y] = P[Y_1 \le y, Y_2 \le y, \ldots, Y_n \le y]$$

$$= \left(\frac{y}{\theta}\right)^n$$

$$\Rightarrow f_{Y_{(n)}}(y) = n\left(\frac{y}{\theta}\right)^{n-1}\cdot\frac{1}{\theta} = \frac{n y^{n-1}}{\theta^n}, \quad 0 < y < \theta$$

$$\Rightarrow E(Y_{(n)}) = \int_0^\theta \frac{n\cdot y^n}{\theta^n}\,dy = \frac{n\cdot y^{n+1}}{(n+1)\cdot\theta^n}\Bigg|_0^\theta = \frac{n\cdot\theta^{n+1}}{(n+1)\cdot\theta^n}$$

$$= \frac{n\theta}{n+1}$$

$$\Rightarrow E(\hat{\theta}_2) = E\left[\frac{n+1}{n} Y_{(n)}\right] = \frac{n+1}{n}\cdot\frac{n\theta}{n+1} = \boxed{\theta}$$ (i.e. $\hat{\theta}_2$ is unbiased)

$$E(Y_{(n)}^2) = \int_0^\theta \frac{n\cdot y^{n+1}}{\theta^n}\,dy = \frac{n\cdot y^{n+2}}{(n+2)\theta^n}\Bigg|_0^\theta = \frac{n\theta^2}{(n+2)}$$

$$\Rightarrow E(\hat{\theta}_2^2) = E\left[\left(\frac{n+1}{n} Y_{(n)}\right)^2\right] = \left(\frac{n+1}{n}\right)^2 \cdot \frac{n\theta^2}{n+2}$$

$$\Rightarrow var(\hat{\theta}_2) = \frac{(n+1)^2\theta^2}{n(n+2)} - \theta^2$$

$$= \frac{(n^2+2n+1)\theta^2 - (n^2+2n)\theta^2}{n(n+2)}$$

$$= \frac{\theta^2}{n(n+2)}$$

$$\Rightarrow eff(\hat{\theta}_1, \hat{\theta}_2) = \frac{\frac{\theta^2}{n(n+2)}}{\frac{\theta^2}{3n}} = \frac{3}{n+2} < 1 \text{ for } n > 1$$

$\Rightarrow \hat{\theta}_2 = \frac{n+1}{n} Y_{(n)}$ is better than $\hat{\theta}_1 = 2\bar{Y}$

e.g. 9.5 let $y_1 \dots y_n$ be a r.s. from $Y \sim N(\mu, \sigma^2)$

$$\hat{\sigma}_1^2 = S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2 \qquad \hat{\sigma}_2^2 = \frac{1}{2}(Y_1 - Y_2)^2$$

Both estimators are unbiased, find eff $(\hat{\sigma}_1, \hat{\sigma}_2)$

$$W = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)} \Rightarrow E(W) = n-1 \quad var(W) = 2(n-1)$$

$$S^2 = \frac{\sigma^2 W}{n-1} \Rightarrow var(S^2) = \frac{\sigma^4}{(n-1)^2} \cdot var(W) = \boxed{\frac{2\sigma^4}{n-1}}$$

$$var(\hat{\sigma}_2^2) = var\left[\frac{1}{2}(Y_1 - Y_2)^2\right]$$

$$Y_1 - Y_2 \sim N(\mu_{Y_1-Y_2} = 0,\ \sigma^2_{Y_1-Y_2} = 2\sigma^2)$$

$$\Rightarrow \frac{Y_1 - Y_2}{\sqrt{2}\,\sigma} \sim N(0,1)$$

$$\Rightarrow \frac{1}{2\sigma^2}(Y_1 - Y_2)^2 \sim \chi^2_{(1)}$$

$$\Rightarrow var\left[\frac{1}{2\sigma^2}(Y_1 - Y_2)^2\right] = 2$$

$$\Rightarrow var(\hat{\sigma}_2^2) = var\left[\sigma^2\left\{\frac{1}{2\sigma^2}(Y_1 - Y_2)^2\right\}\right] = \sigma^4 \cdot 2$$

$$\Rightarrow eff(\hat{\theta}_1, \hat{\theta}_2) = \frac{2\sigma^4}{\frac{2\sigma^4}{(n-1)}} = n-1 > 1 \quad \text{for } n > 1$$

$\Rightarrow$ $\hat{\theta}_1$ is better than $\hat{\theta}_2$

9.8 <u>Cramer-Rao Lower Bound</u>

$Y_1 \cdots Y_n$ is a r.s. from $f(y)$ with unknown parameter θ. If $\hat{\theta}$ is an unbiased estimator of θ then under very general conditions...

$$\boxed{V(\hat{\theta}) \geq \frac{1}{nE\left[-\frac{\partial^2 \ln f(Y)}{\partial \theta^2}\right]} = I(\hat{\theta})}$$

<u>note:</u> If $v(\hat{\theta}) = \dfrac{1}{nE\left[\frac{-\partial^2 \ln f(Y)}{\partial \theta^2}\right]}$ then $\hat{\theta}$ is

called a uniform minimum variance unbiased estimator or UMVUE for Θ.

Let $f(y)$ have normal density w/ mean $= \mu$ and variance $= \sigma^2$. Show $\bar{Y}$ is UMVUE for μ.

$$\ln f(y) = \ln \left[\frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2\sigma^2}(y-\mu)^2} \right]$$

$$= -\tfrac{1}{2}\ln(2\pi) - \ln\sigma - \frac{1}{2\sigma^2}(y-\mu)^2$$

$$\frac{\partial \ln f(y)}{\partial \mu} = -\frac{1}{2\sigma^2} \cdot 2(y-\mu)(-1) = \frac{y-\mu}{\sigma^2}$$

$$\frac{\partial^2 \ln f(y)}{\partial \mu^2} = -\frac{1}{\sigma^2}$$

$$E\left[-\frac{\partial^2 \ln f(y)}{\partial \mu^2}\right] = E\left[\frac{1}{\sigma^2}\right] = \frac{1}{\sigma^2}$$

$$\Rightarrow I(\hat{\theta}) = \frac{1}{n\left(\frac{1}{\sigma^2}\right)} = \frac{\sigma^2}{n}$$

Since $V(\bar{Y}) = \frac{\sigma^2}{n} = I(\hat{\theta})$ (or CRLB)

$\Rightarrow \bar{Y}$ is UMVUE for μ ∎

9.3 Consistency

Types of convergence

Def Convergence in distribution

Let $\{Y_n\}$ be a sequence of RVs and let Y be a RV. Let $C[F(y)]$ denote the set of all pts where $F(y)$ is cont.

We say that Y_n converges in dist. to Y if

$$\lim_{n\to\infty} F_{Y_n}(y) = F(y) \quad \forall y \in C[F(y)]$$

denoted by $Y_n \xrightarrow{D} Y$

Def Convergence in probability

Let $\{Y_n\}$ be a sequence of RVs and let Y be a RV defined on a sample space. We say Y_n converges in probability to Y if $\forall \varepsilon > 0$

$$\lim_{n\to\infty} P[|Y_n - Y| \le \varepsilon] = 1$$

denoted by $Y_n \xrightarrow{P} Y$

Def 9.2 The estimator $\hat{\theta}_n$ is said to be a consistent estimator of θ if, for any positive no. ε,

$$\lim_{n\to\infty} P(|\hat{\theta}_n - \theta| \le \varepsilon) = 1$$

or equivalently

$$\lim_{n\to\infty} P(|\hat{\theta}_n - \theta| > \varepsilon) = 0$$

note: $\hat{\theta}_n \xrightarrow{P} \theta$ or convergence in prob to a constant (parameter) is consistency

Weak Law of Large Numbers (WLLN)

Let $Y_1 \cdots Y_n$ be a r.s. from $f(y)$ w/ mean $= \mu$, var $= \sigma^2$

Then $\bar{Y}_n \xrightarrow{P} \mu$

Pf:

$E(\bar{Y}_n) = \mu$ and $var(\bar{Y}_n) = \sigma^2/n$

By Chebychev's rule..

$$P\left[|\bar{Y}_n - \mu| \le k\sigma_{\bar{Y}_n} \right] \ge 1 - \frac{1}{k^2}$$

$$P\left[|\bar{Y}_n - \mu| \le k \cdot \sigma/\sqrt{n} \right] \ge 1 - \frac{1}{k^2}$$

let $\varepsilon = k \cdot \sigma/\sqrt{n}$

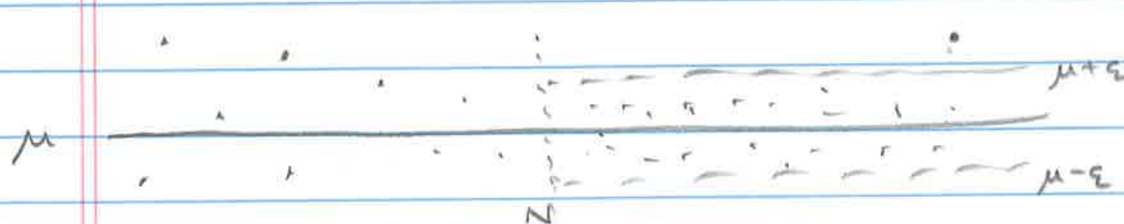
$$P\left[|\bar{Y}_n - \mu| \le \varepsilon \right] \ge 1 - \frac{1}{\left(\frac{\varepsilon \cdot \sqrt{n}}{\sigma}\right)^2} = 1 - \frac{\sigma^2}{n^2 \varepsilon^2}$$

$$\lim_{n \to \infty} P\left[|\bar{Y}_n - \mu| \le \varepsilon \right] \ge \lim_{n \to \infty} \left(1 - \frac{\sigma^2}{n^2 \varepsilon^2}\right) = 1$$

$$\lim_{n \to \infty} P\left[|\bar{Y}_n - \mu| \le \varepsilon \right] = 1 \quad \Rightarrow \quad \bar{Y}_n \xrightarrow{P} \mu$$

i.e. $\bar{Y}_n$ is a consistent estimator of μ

Conceptually ....



For $n > N$, there is a very high probability (not guaranteed) that $\bar{y}$ will be within ε of μ. The farther out you go, the less likely you will find pts. outside the limits, but you can never guarentee all $\bar{y}_n$ will be within ε of μ.

$\boxed{\text{coin toss in excel}}$ $\Rightarrow$ $\hat{p}$ is a consistent estimator of p.

e.g. Let $Y \sim \text{unif}(0, \theta)$

Is $\hat{\theta} = 2\bar{Y}$ consistent? (For unbiased estimators we can use Chebychev's rule)

$$E(\hat{\theta}) = 2E(\bar{Y}) = 2 \cdot \frac{\theta}{2} = \theta \quad \text{(unbiased)}$$

$$V(\hat{\theta}) = 4\,\text{var}(\bar{Y}) = 4 \cdot \frac{\frac{\theta^2}{12}}{n} = \frac{\theta^2}{3n}$$

$$P\left[\,|\hat{\theta} - \theta| \le k \cdot \sigma_{\hat{\theta}}\,\right] \ge 1 - \frac{1}{k^2}$$

$$\text{let } \varepsilon = k \cdot \sigma_{\hat{\theta}} = k \cdot \sqrt{\frac{\theta^2}{3n}} = \frac{k \cdot \theta}{\sqrt{3n}}$$

$$P\left[\,|\hat{\theta} - \theta| \le \varepsilon\,\right] \ge 1 - \frac{1}{\left[\frac{\varepsilon(\sqrt{3n})}{\theta}\right]^2} = 1 - \frac{\theta^2}{\varepsilon^2(3n)}$$

$$\lim_{n \to \infty} P\left[\,|\hat{\theta} - \theta| \le \varepsilon\,\right] \ge \lim_{n \to \infty}\left(1 - \frac{\theta^2}{\varepsilon^2(3n)}\right) = 1$$

$$\Rightarrow \lim_{n \to \infty} P\left[\,|\hat{\theta} - \theta| \le \varepsilon\,\right] = 1 \quad \text{i.e. } \hat{\theta} = 2\bar{Y} \text{ is consistent for } \theta$$

<u>Thm 9.1</u> an unbiased estimator $\hat{\theta}_n$ is a consistent estimator of θ if

$$\lim_{n\to\infty} V(\hat{\theta}_n) = 0$$

<u>pf:</u> again we will use Chebychev's Rule

$$P\left[|\hat{\theta}_n - \theta| \le k\sigma_{\hat{\theta}}\right] \ge 1 - \frac{1}{k^2}$$

let $\varepsilon = k\sigma_{\hat{\theta}}$

$$P\left[|\hat{\theta}_n - \theta| \le \varepsilon\right] \ge 1 - \frac{1}{(\varepsilon/\sigma_{\hat{\theta}})^2} = 1 - \left(\frac{\sigma_{\hat{\theta}}}{\varepsilon}\right)^2$$

$$\lim_{n\to\infty}\left(1 - \frac{\sigma^2_{\hat{\theta}}}{\varepsilon^2}\right) = 1 - \frac{\lim_{n\to\infty}\sigma^2_{\hat{\theta}} \to 0}{\varepsilon^2} = 1$$

$$\Rightarrow \lim_{n\to\infty} P\left[|\hat{\theta}_n - \theta| \le \varepsilon\right] \ge \lim_{n\to\infty}\left(1 - \frac{\sigma^2_{\hat{\theta}}}{\varepsilon^2}\right) = 1$$

$$\Rightarrow \lim_{n\to\infty} P\left[|\hat{\theta}_n - \theta| \le \varepsilon\right] = 1 \quad \Rightarrow \hat{\theta} \text{ consistent for } \theta$$

<u>note:</u> WLLN : $E(\bar{Y}_n) = \mu \quad V(\bar{Y}_n) = \sigma^2/n$

$\lim_{n\to\infty} \sigma^2/n = 0 \Rightarrow \bar{Y}_n$ is consistent for μ

also, $\lim_{n\to\infty} \frac{p(1-p)}{n} = 0$

$\Rightarrow \hat{p}$ <u>consistent</u> for p

e.g.: $Y_1 \cdots Y_n$ is a r.s. from $Y \sim \text{unif}(0, \theta)$

$\hat{\theta}_1 = 2\bar{Y} \qquad \hat{\theta}_2 = \frac{n+1}{n} Y_{(n)}$

We showed in 9.2 that $E(\hat{\theta}_1) = E(\hat{\theta}_2) = \theta$ (unbiased) also,

$$V(\hat{\theta}_1) = \frac{\theta^2}{3n} \qquad \lim_{n\to\infty} \frac{\theta^2}{3n} = 0 \Rightarrow \hat{\theta}_1 \text{ consistent}$$

$$V(\hat{\theta}_2) = \frac{\theta^2}{n(n+2)} \qquad \lim_{n\to\infty} \frac{\theta^2}{n(n+2)} = 0 \Rightarrow \hat{\theta}_2 \text{ consistent}$$

<u>note:</u> $\hat{\theta}_2$ was more efficient than $\hat{\theta}_1$

9.21 Let $Y_1 \ldots Y_n$ be a r.s. from $Y \sim N(\mu, \sigma)$

assume $n = 2k$ (k an integer)

$$\hat{\sigma}^2 = \frac{1}{2k} \sum_{i=1}^{k} (Y_{2i} - Y_{2i-1})^2$$

a) $$E\left[\frac{1}{2k} \sum_{i=1}^{k} (Y_{2i} - Y_{2i-1})^2\right]$$

$$= \frac{1}{2k} \sum_{i=1}^{k} E\left[(Y_{2i} - Y_{2i-1})^2\right]$$

$$= \frac{1}{2k} \sum_{i=1}^{k} E\left[Y_{2i}^2 - 2Y_{2i} \cdot Y_{2i-1} + Y_{2i-1}^2\right]$$

$$= \frac{1}{2k} \sum_{i=1}^{k} \left[E(Y_{2i}^2) - 2E(Y_{2i}) \cdot E(Y_{2i-1}) + E(Y_{2i-1}^2)\right]$$

$$= \frac{1}{2k} \cdot k\left[(\sigma^2 + \mu^2) - 2\mu^2 + (\sigma^2 + \mu^2)\right]$$

$$= \frac{1}{2}\left[\sigma^2 + \sigma^2\right] = \frac{2\sigma^2}{2} = \sigma^2 \Rightarrow \text{unbiased}$$

b) $Y_{2i} - Y_{2i-1} \sim N(\mu = 0, \text{var} = 2\sigma^2)$

$$\Rightarrow \frac{Y_{2i} - Y_{2i-1}}{\sqrt{2}\,\sigma} \sim N(0,1)$$

$$\Rightarrow \frac{(Y_{2i}-Y_{2i-1})^2}{2\sigma^2} \sim \chi^2_{(1)}$$

$$\Rightarrow \sum_{i=1}^{k} \frac{(Y_{2i}-Y_{2i-1})^2}{2\sigma^2} \sim \chi^2_{(k)}$$

Let $W = \sum_{i=1}^{k} \frac{(Y_{2i}-Y_{2i-1})^2}{2\sigma^2}$ $\Rightarrow V(W) = 2k$

$$\Rightarrow \hat{\sigma}^2 = \frac{2\sigma^2 \cdot W}{2k} = \frac{\sigma^2 W}{k}$$

$$\Rightarrow var(\hat{\sigma}^2) = \frac{\sigma^4}{k^2} \cdot var(W) = \frac{\sigma^4 (2k)}{k^2} = \frac{2\sigma^4}{k} = \frac{4\sigma^4}{n}$$

$$\Rightarrow \lim_{n\to\infty} \frac{4\sigma^4}{n} = 0 \Rightarrow \hat{\sigma}^2 \text{ is consistent for } \sigma^2$$

(Thm 9.1)

Thm 9.2

Spse that $\hat{\theta}_n \xrightarrow{P} \theta$ and $\hat{\theta}'_n \xrightarrow{P} \theta'$ then

a) $\hat{\theta}_n + \theta_n \xrightarrow{P} \theta + \theta'$

b) $\hat{\theta}_n \times \hat{\theta}'_n \xrightarrow{P} \theta \times \theta'$

c) If $\theta' \neq 0$, $\hat{\theta}_n / \hat{\theta}'_n \xrightarrow{P} \theta / \theta'$

d) If $g(\cdot)$ is a real-valued fct that is cont. at θ, then $g(\hat{\theta}_n) \xrightarrow{P} g(\theta)$

Pf: (d)

Because g is cont., it follows that for every $\varepsilon > 0$, a $\delta > 0$ exists such that $|x-\theta| < \delta \Rightarrow |g(x) - g(\theta)| < \varepsilon$

since $|x-\theta|<\delta \Rightarrow |g(x)-g(\theta)|<\varepsilon$, we can say that the "event" $|x-\theta|<\delta$ is "contained" in the "event" $|g(x)-g(\theta)|<\varepsilon$.

$$\Rightarrow P(|g(x)-g(\theta)|<\varepsilon) \geq P(|x-\theta|<\delta)$$

substitute $\hat{\theta}_n$ in for x and ...

$$P(|g(\hat{\theta}_n)-g(\theta)|<\varepsilon) \geq P(|\hat{\theta}_n-\theta|<\delta)$$

It is given that $\lim\limits_{n\to\infty} P(|\hat{\theta}_n-\theta|<\delta)=1$

$$\Rightarrow \lim_{n\to\infty} P(|g(\hat{\theta}_n)-g(\theta)|<\varepsilon)=1$$

$$\Rightarrow g(\hat{\theta}_n) \xrightarrow{P} g(\theta)$$

e.g. Spse $Y\sim Bin(n,p)$ Let $\hat{p}_n = Y/n$

Show that $\hat{p}_n(1-\hat{p}_n) \xrightarrow{P} p(1-p)$

$E(\hat{p}_n)=p \quad var(\hat{p}_n)=\frac{p(1-p)}{n} \quad \lim\limits_{n\to\infty} \frac{p(1-p)}{n}=0$

$\Rightarrow \hat{p}_n$ is consistent for p

$\Rightarrow g(\hat{p}_n) \xrightarrow{P} g(p)$ let $g(\hat{p}_n)=\hat{p}_n(1-\hat{p}_n) \Rightarrow g(p)=p(1-p)$

$\Rightarrow \hat{p}_n(1-\hat{p}_n) \xrightarrow{P} p(1-p)$ ∎

$\boxed{\text{Example 9.3}}$

Show that $S_n^2 \xrightarrow{P} \sigma^2$ if $E(Y_i)=\mu$, $E(Y_i^2)=\mu_2'$, $E(Y_i^4)=\mu_4'$ all finite

$$S_n^2=\frac{1}{n-1}\sum(Y_i-\bar{Y})^2=\frac{1}{n-1}\left[\sum Y_i^2-n\bar{Y}^2\right]=\frac{n}{n-1}\left[\frac{1}{n}\sum Y_i^2-\bar{Y}^2\right]$$

$\bar{Y} \xrightarrow{P} \mu$ so by Thm 9.2(d) $\bar{Y}^2 \xrightarrow{P} \mu^2$ (note: $g(x)=x^2$ is cont. $-\infty<x<\infty$)

$var(Y_i^2) = E(Y_i^4) - E(Y_i^2)^2 = \mu_4' - \mu_2'^2 < \infty$ which is finite

$var\left(\frac{1}{n}\sum Y_i^2\right) \underset{\uparrow}{=} \frac{1}{n^2}\sum var(Y_i^2) = \frac{1}{n^2}\left(n\left(\mu_4' - \mu_2'^2\right)\right)$

Ind. w/ r.s.

$$= \frac{\mu_4' - \mu_2'^2}{n}$$

$\lim_{n\to\infty} var\left(\frac{1}{n}\sum Y_i^2\right) = \lim_{n\to\infty} \frac{\mu_4' - \mu_2'^2}{n} = 0$

$E\left[\frac{1}{n}\sum Y_i^2\right] = \frac{1}{n}\sum E(Y_i^2) = \frac{n\cdot\mu_2'}{n} = \mu_2'$ (unbiased)

$\Rightarrow$ by Thm 9.1 $\frac{1}{n}\sum Y_i^2 \xrightarrow{P} \mu_2'$

$\Rightarrow$ by Thm 9.2(a) $\frac{1}{n}\sum Y_i^2 - \bar{Y}^2 \xrightarrow{P} \mu_2' - \mu^2 = \sigma^2$

$\frac{n}{n-1}$ is a series of constants that converge to 1

let $W_n = \frac{n}{n-1}$ then $P\left[|W_n - 1| < \varepsilon\right] = P\left[\left|\frac{n}{n-1} - 1\right| < \varepsilon\right]$

$= P\left[-\varepsilon < \frac{n}{n-1} - 1 < \varepsilon\right] = P\left[-\varepsilon < \frac{1}{n-1} < \varepsilon\right]$

$\overset{n>1}{=} P\left[\frac{1}{n-1} < \varepsilon\right] = 1$ for $n > \frac{1+\varepsilon}{\varepsilon}$

$\Rightarrow \lim_{n\to\infty} P\left[|W_n - 1| < \varepsilon\right] = 1 \Rightarrow W_n \xrightarrow{P} 1$

by Thm 9.2(b) $\frac{n}{n-1}\left[\frac{1}{n}\sum Y_i^2 - \bar{Y}^2\right] \xrightarrow{P} 1\cdot\sigma^2 = \sigma^2$ ∎

Thm 9.3

Spse that U_n has a dist. fct. that converges to a $N(0,1)$ dist. fct. as $n \to \infty$ (i.e. $U_n \xrightarrow{D} Z$). If W_n converges in prob. to 1 (i.e. $W_n \xrightarrow{P} 1$) then the dist. fct. of U_n / W_n converges to a $N(0,1)$ (i.e. $U_n/W_n \xrightarrow{D} Z$)

Example 9.4

Spse $Y_1 \cdots Y_n$ is a r.s. of n from $f(y)$ with μ, σ^2

Let $S_n = \frac{1}{n-1}(Y_i - \bar{Y}_n)^2$

Show $\sqrt{n}\left(\frac{\bar{Y}_n - \mu}{S_n}\right) \xrightarrow{d} N(0,1)$

$$\sqrt{n}\left(\frac{\bar{Y}_n - \mu}{S_n}\right) = \frac{\bar{Y}_n - \mu}{S_n/\sqrt{n}} = \frac{\bar{Y}_n - \mu}{\sigma/\sqrt{n}} \cdot \frac{1}{(S_n/\sigma)}$$

$\frac{\bar{Y}_n - \mu}{\sigma/\sqrt{n}} \xrightarrow{d} N(0,1)$ (By CLT in 7.4)

$S_n \xrightarrow{P} \sigma$ (Ex 9.3). Let $g(x) = \sqrt{\frac{x}{c}}$ then g is cont.

for $x > 0$, $c > 0$

$$\Rightarrow \frac{S_n}{\sigma} = \sqrt{\frac{S_n^2}{\sigma^2}} \xrightarrow{P} \sqrt{\frac{\sigma^2}{\sigma^2}} = 1$$

$$\Rightarrow \text{by Thm 9.3} \quad \sqrt{n}\left(\frac{\bar{Y}_n - \mu}{S_n}\right) = \frac{\bar{Y}_n - \mu}{\sigma/\sqrt{n}} \cdot \left(\frac{1}{S_n/\sigma}\right) \xrightarrow{d} N(0,1)$$

what is the significance of this result? It justifies the large sample CI for μ: $(\bar{y} \pm z_{\alpha/2} \cdot s/\sqrt{n})$
(recall simulation)

note:

① This result is valid regardless of the shape of the dist. we sample from

② If $Y_1 \cdots Y_n$ is a r.s. from $Y \sim N(\mu, \sigma)$ then

$$\frac{\bar{Y}_n - \mu}{s_n/\sqrt{n}} \sim t_{(n-1)}$$

Therefore, this shows that $t \xrightarrow{d} z$ (This is why our t-tables stop for large no. of d.f.)

9.36 $Y \sim Bin(n, p)$

let $\hat{p}_n = \frac{Y}{n}$ show $\frac{\hat{p}_n - p}{\sqrt{\frac{\hat{p}_n \hat{q}_n}{n}}} \xrightarrow{d} N(0,1)$

From ch. 7. (CLT)

$$\frac{Y - np}{\sqrt{np(1-p)}} = \frac{\hat{p}_n - p}{\sqrt{\frac{p(1-p)}{n}}} \xrightarrow{d} N(0,1)$$

Earlier we showed $\hat{p}_n(1-\hat{p}_n) \xrightarrow{P} p(1-p)$

let $g(x) = \sqrt{\frac{x}{c}}$, g is cont. for $x > 0, c > 0$

$$\Rightarrow g(\hat{p}_n(1-\hat{p}_n)) = \sqrt{\frac{\hat{p}_n(1-\hat{p}_n)}{p(1-p)}} \xrightarrow{P} \sqrt{\frac{p(1-p)}{p(1-p)}} = 1$$

By Thm 9.3 ...

$$\frac{\dfrac{\hat{p}_n - p}{\sqrt{\frac{p(1-p)}{n}}}}{\sqrt{\dfrac{\hat{p}_n(1-\hat{p}_n)}{p(1-p)}}} \xrightarrow{d} N(0,1)$$

$$= \frac{\hat{p}_n - p}{\sqrt{\frac{p(1-p)}{n}}} \cdot \frac{\sqrt{p(1-p)}}{\sqrt{\hat{p}_n(1-\hat{p}_n)}} = \frac{\hat{p}_n - p}{\sqrt{\frac{\hat{p}_n(1-\hat{p}_n)}{n}}}$$

This justifies large sample CI for p: $\boxed{\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}(\hat{q})}{n}}}$

9.28 $Y_1 \cdots Y_n$ is a r.s. from a pareto dist.

$$\Rightarrow F_{(1)}(y) = \begin{cases} 0 & , y \le \beta \\ 1-(\beta/y)^{\alpha n}, & y > \beta \end{cases}$$

Show $Y_{(1)} \xrightarrow{p} \beta$ using the definition of convergence in prob.

$$P[|Y_{(1)} - \beta| < \varepsilon] = P[-\varepsilon < Y_{(1)} - \beta < \varepsilon] = P[\beta - \varepsilon < Y_{(1)} < \beta + \varepsilon]$$

$$= P(Y_{(1)} < \beta+\varepsilon) = 1 - \left(\beta/\beta+\varepsilon\right)^{\alpha n}$$

$$\lim_{n\to\infty} P[|Y_{(1)} - \beta| < \varepsilon] = \lim_{n\to\infty}\left[1 - \left(\frac{\beta}{\beta+\varepsilon}\right)^{\alpha n}\right] = 1 - \lim_{n\to\infty}\left\{\left(\frac{\beta}{\beta+\varepsilon}\right)^{\alpha}\right\}^{n}$$

$$0 < \left(\frac{\beta}{\beta+\varepsilon}\right)^{\alpha} < 1 \text{ for } \varepsilon > 0 \Rightarrow \lim_{n\to\infty}\left\{\left(\frac{\beta}{\beta+\varepsilon}\right)^{\alpha}\right\}^{n} = 0$$

$$\Rightarrow \lim_{n\to\infty} P(|Y_{(1)} - \beta| < \varepsilon) = 1 - 0 = 1 \Rightarrow Y_{(1)} \xrightarrow{p} \beta \blacksquare$$

9.21

Y_i iid $\sim N(\mu, \sigma^2)$

$n = 2k$

$$\hat{\sigma}^2 = \frac{1}{2k} \sum_{i=1}^{k} (Y_{2i} - Y_{2i-1})^2$$

$$= \frac{1}{2k} \sum_{i=1}^{k} (Y_{2i}^2 - 2 Y_{2i} Y_{2i-1} + Y_{2i-1}^2)$$

$$E(\hat{\sigma}^2) = \frac{1}{2k} \sum_{i=1}^{k} \left((\mu^2+\sigma^2) - 2\mu^2 + (\mu^2+\sigma^2) \right)$$

$$= \frac{1}{2k} (2k\sigma^2) = \sigma^2$$

$$\mathrm{var}(Y_{2i}^2) = E(Y_{2i}^4) - E(Y_{2i}^2)^2$$

$$= \mu_4' - (\mu^2+\sigma^2)^2$$

$$\mathrm{var}(Y_{2i} \cdot Y_{2i-1}) = E(Y_{2i}^2 \cdot Y_{2i-1}^2) - E(Y_{2i} \cdot Y_{2i-1})^2$$

$$= (\mu^2+\sigma^2)^2 - \mu^4$$

$$\mathrm{var}(\hat{\sigma}^2) = \frac{1}{(2k)^2} \left[k(\mu_4' - (\mu^2+\sigma^2)^2) + 4k((\mu^2+\sigma^2)^2 - \mu^4) + k(\mu_4' - (\mu^2+\sigma^2)^2) \right]$$

$$= \frac{1}{(2k)^2} \left[2k \cdot \mu_4' - 2k(\mu^2+\sigma^2)^2 + 4k(\mu^2+\sigma^2)^2 - 4k \cdot \mu^4 \right]$$

$$= \frac{2k \cdot \mu_4' + 2k(\mu^2+\sigma^2)^2 - 4k\mu^4}{(2k)^2}$$

$$= \frac{\mu_4' + (\mu^2+\sigma^2)^2}{\underbrace{2k}_{\to n}} - \frac{\mu^4}{\underbrace{k}_{\to n/2}}$$

$$= \frac{\mu_4' + (\mu^2+\sigma^2)^2}{n} - \frac{2\mu^4}{n}$$

$$\lim_{n\to\infty} \mathrm{var}(\hat{\sigma}^2) = 0$$

$\Rightarrow \hat{\sigma}^2$ consistent for σ^2

9.4 Sufficiency

Goal: In some cases, it is possible to show that a particular statistic or set of statistics contains all the "info" in the sample about the parameters. It would then be reasonable to restrict our attention to such statistics when estimating (and making inferences) about the parameters.

More generally, the idea of sufficiency involves the reduction of the data set to a more precise set of statistics with no loss of information about the unknown parameter. (i.e. once the value of a sufficient statistic is known, the observed value of any other statistic does not contain any further info about the parameter.

e.g. Let $Y_1 \cdots Y_n$ be a r.s. from $Y \sim POI(\lambda)$

$$\Rightarrow p(y) = \frac{e^{-\lambda}\lambda^y}{y!} \quad , y = 0, 1, 2, \ldots$$

Let $U = \sum_{i=1}^{n} Y_i \sim POI(n\lambda)$ (see 6.5)

$$P(Y_1 = y_1, \ldots, Y_n = y_n \mid U = u) = \frac{P[Y_1 = y_1, \ldots, Y_n = y_n, U = u]}{P[U = u]}$$

note: The numerator is 0 if $\sum y_i \neq u$. It is actually the prob of any sample where $Y_1 = y_1$ and $Y_2 = y_2$ and $\cdots$ and $Y_n = y_n$ and $\sum Y_i = u$.

Therefore it reduces to $P[Y_1 = y_1, \ldots, Y_n = y_n] \stackrel{\perp}{=} \prod_{i=1}^{n} \frac{e^{-\lambda}\lambda^{y_i}}{y_i!}$

for any case where $\sum Y_i = u$.

The denominator only requires that $\Sigma Y_i = u$. It places no restrictions on $Y_1, Y_2, \dots, Y_n$ other than that they sum to u.

$$P[U=u] = \frac{e^{-n\lambda}(n\lambda)^u}{u!} = \frac{e^{-n\lambda} n^{\Sigma Y_i} \cdot \lambda^{\Sigma Y_i}}{(\Sigma Y_i)!}$$

---

e.g. $n=2$ $\quad P[Y_1 = 1, Y_2 = 2 \mid U = 3]$

$$= \frac{P[Y_1 = 1, Y_2 = 2, U=3]}{P[U=3]}$$

redundant

4 cases

Y_1	Y_2
0	3
1	2
2	1
0	3

$$\Rightarrow P[Y_1 = y_1, \dots, Y_n = y_n \mid U = u] = \frac{\prod_{i=1}^{n} \frac{e^{-\lambda}\lambda^{y_i}}{y_i!}}{\frac{e^{-n\lambda} n^{\Sigma y_i} \lambda^{\Sigma y_i}}{(\Sigma y_i)!}} = \frac{\frac{e^{-n\lambda}\lambda^{\Sigma y_i}}{\prod_{i=1}^{n} y_i!}}{\frac{e^{-n\lambda} n^{\Sigma y_i} \lambda^{\Sigma y_i}}{(\Sigma y_i)!}}$$

$$= \frac{(\Sigma y_i)!}{\prod_{i=1}^{n} y_i! \cdot n^{\Sigma y_i}}$$

note: This does not depend on θ. Thus it is sufficient to examine Σx_i to estimate θ. All other info is not needed for θ.

i.e. Any fct of Σy_i considered to be a suff. stat.

Def 9.3 Let $Y_1 \cdots Y_n$ denote a r.s. from a prob. dist. with unknown parameter θ. Then the statistic $U = g(Y_1 \cdots Y_n)$

is said to be sufficient for θ iff.

$$f(y_1,\ldots,y_n|u) = \frac{\prod_{i=1}^{n} f(y_i)}{f(u)} = H(y_1,\ldots,y_n)$$

where $H(y_1,\ldots,y_n)$ does not depend on θ for every fixed value of $u = g(y_1,\ldots,y_n)$

i.e. neither the formula, nor the domain of $f(y_1,\ldots y_n|u)$ can contain θ. (note: if support for $f(y)$ depends on θ, use indicator fcts)

e.g. Let $Y_1 \ldots Y_n$ be a r.s. from $Y \sim N(\mu, \sigma^2)$. Show $\bar{Y}$ and S^2 are jointly suff. for μ, σ^2

note: $\bar{Y}$ and S^2 are ind. when sampling from a normal dist.

$$\bar{Y} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \qquad \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$$

Let $W = \frac{(n-1)S^2}{\sigma^2}$ $\quad m_W(t) = \frac{1}{(1-2t)^{(n-1)/2}}$

$$S^2 = \frac{\sigma^2 \cdot W}{n-1} \qquad m_{S^2}(t) = E\left[e^{tS^2}\right] = E\left[e^{t\left(\frac{\sigma^2 W}{n-1}\right)}\right]$$

$$= m_W\left(\frac{\sigma^2 t}{n-1}\right) = \frac{1}{\left(1 - \frac{2\sigma^2}{n-1}t\right)^{\frac{n-1}{2}}} \sim GAM\left(\alpha = \frac{n-1}{2}, \beta = \frac{2\sigma^2}{n-1}\right)$$

$$f(y_1,\ldots,y_n \mid \bar{y}, s^2) = \frac{f(y_1,\ldots,y_n, \bar{y}, s^2)}{f(\bar{y}, s^2)}$$

redundant

$$= \frac{f(y_1,\ldots,y_n)}{f(\bar{y}, s^2)} = \frac{\prod_{i=1}^{n} f(y_i)}{f(\bar{y}) \cdot f(s^2)}$$

$$= \frac{\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\frac{1}{2\sigma^2}\Sigma(y_i-\mu)^2}}{\left[\frac{1}{\sqrt{2\pi\sigma^2/n}} e^{-\frac{n}{2\sigma^2}(\bar{y}-\mu)^2}\right]\left[\frac{1}{\Gamma(\frac{n-1}{2})\left(\frac{2\sigma^2}{n-1}\right)^{\frac{n-1}{2}}}(s^2)^{\frac{n-1}{2}-1}\cdot e^{-\frac{s^2}{(2\sigma^2/(n-1))}}\right]}$$

$$= \frac{\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\frac{1}{2\sigma^2}\left[\Sigma(y_i-\bar{y})^2+n(\bar{y}-\mu)^2\right]}}{\left[\frac{1}{\sqrt{2\pi\sigma^2/n}} e^{-\frac{n}{2\sigma^2}(\bar{y}-\mu)^2}\right]\left[\frac{1}{\Gamma(\frac{n-1}{2})\left(\frac{2\sigma^2}{n-1}\right)^{\frac{n-1}{2}}}(s^2)^{\frac{n-1}{2}-1} e^{-\frac{n-1}{2\sigma^2}\left(\frac{\Sigma(y_i-\bar{y})^2}{n-1}\right)}\right]}$$

$$= \frac{\frac{1}{(2\pi\sigma^2)^{n/2}}\cdot \cancel{e^{-\frac{1}{2\sigma^2}\Sigma(y_i-\bar{y})^2}}\cdot \cancel{e^{-\frac{n}{2\sigma^2}(\bar{y}-\mu)^2}}}{\frac{1}{\left(\frac{2\pi}{n}\sigma^2\right)^{1/2}}\cdot\frac{1}{\Gamma(\frac{n-1}{2})\left(\frac{2\cdot\sigma^2}{n-1}\right)^{\frac{n-1}{2}}}(s^2)^{\frac{n-1}{2}-1}\cdot \cancel{e^{-\frac{n}{2\sigma^2}(\bar{y}-\mu)^2}}\cdot \cancel{e^{-\frac{1}{2\sigma^2}\Sigma(y_i-\bar{y})^2}}}$$

$$= \frac{\frac{1}{k_1\sigma^n}}{\frac{1}{k_2\cdot\sigma}\cdot\frac{1}{k_3\cdot\sigma^{n-1}}\cdot(s^2)^{\frac{n-1}{2}-1}} = \frac{k}{(s^2)^{\frac{n-1}{2}-1}}$$

whew!! does not depend on μ, σ^2 $\Rightarrow$ $\bar{y}$ and s^2 are jointly sufficient for μ, σ^2.

note: (1) Difficult using definition

(2) Definition only allows us to check for sufficient

statistics. It does not help us find them.

Def 9.4 Let $y_1 \cdots y_n$ be sample observations taken on corresponding RVs $Y_1 \cdots Y_n$ whose dist. depends on parameter θ.

a) If $Y_1 \cdots Y_n$ are discrete, the likelihood of the sample is

$$L(y_1, y_2, \ldots, y_n \mid \theta) = p(y_1, \ldots, y_n \mid \theta)$$

or

$$L(\theta) = \boxed{\prod_{i=1}^{n} p(y_i \mid \theta)}$$

b) If $Y_1 \cdots Y_n$ are cont. RVs, the likelihood of the sample is

$$L(y_1, \ldots, y_n \mid \theta) = f(y_1, \ldots, y_n \mid \theta)$$

or

$$L(\theta) = \boxed{\prod_{i=1}^{n} f(y_i \mid \theta)}$$

Thm 9.4 Let u be a statistic based on the r.s. $Y_1 \cdots Y_n$. Then U is a sufficient statistic for the estimation of a parameter θ iff. $L(\theta)$ can be factored into two non-negative fcts,

$$\boxed{L(\theta) = g(u, \theta)\, h(y_1, \ldots, y_n)}$$

where $g(u, \theta)$ is a fct. only of u and θ and for every fixed value of u, $h(y_1, \ldots, y_n)$ does not depend on θ.

This is known as the Factorization Thm

Proof:

($\Leftarrow$) assume Y is cont. and $f(y_1, \cdots y_n \mid \theta) = L(\theta) = g(u,\theta) \cdot h(y_1 \cdots y_n)$

Find $f(u)$ using a pdf to pdf transformation:

$u = g(y_1, \dots, y_n)$ inverse fcts: $Y_1 = w_1(u, u_2, \dots, u_n)$

$u_2 = g_2(y_1, \dots, y_n)$ $Y_2 = w_2(u, u_2, \dots, u_n)$

$\vdots$ $\vdots$

$u_n = g_n(y_1, \dots, y_n)$ $Y_n = w_n(u, u_2, \dots, u_n)$

choose $u_2 \cdots u_n$ so that the transformation is 1-1

$$J = \det \begin{bmatrix} \frac{\partial Y_1}{\partial u} & \frac{\partial Y_1}{\partial u_2} & \cdots & \frac{\partial Y_1}{\partial u_n} \\ \vdots & & & \\ \frac{\partial Y_n}{\partial u} & \frac{\partial Y_n}{\partial u_2} & \cdots & \frac{\partial Y_n}{\partial u_n} \end{bmatrix} \quad \text{not a fct of } \theta$$

$\Rightarrow$ joint pdf of $u, u_2, \dots, u_n$ is:

$$f(u, u_2, \dots, u_n) = f(w_1, w_2, \dots, w_n) \cdot |J|$$

$$\overset{\text{given}}{=} g(u,\theta) \cdot h(w_1, \dots, w_n) \cdot |J|$$

$$\Rightarrow f_u(u) = g(u,\theta) \cdot \underbrace{\underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{\text{no } \theta \text{ in limits}} \underset{\text{no } \theta}{h(w_1, \dots, w_n)} \cdot \underset{\text{no } \theta}{|J|} \, du_2 \cdots du_n}_{\text{only a fct. of } u}$$

$$= g(u,\theta) \cdot m(u)$$

$$\Rightarrow f(y_1, \cdots y_n \mid u) = \frac{f(y_1, \dots, y_n)}{f_u(u)} = \frac{\cancel{g(u,\theta)} \cdot h(y_1 \cdots y_n)}{\cancel{g(u,\theta)} \cdot m(u)} = Q(y_1, \dots, y_n)$$

$\Rightarrow u$ is suff for θ ∎

Proof ($\Longrightarrow$)

If u is suff. for θ then

$$f(y_1,\dots,y_n;\theta \mid u) = \frac{f(y_1,\dots,y_n;\theta)}{f_u(u;\theta)} = \underbrace{H(y_1,\dots,y_n)}_{\text{does not depend on } \theta}$$

$$\Rightarrow f(y_1,\dots,y_n;\theta) = f_u(u;\theta)\cdot H(y_1,\dots,y_n)$$

Let $g(u,\theta) = f_u(u;\theta)$

$\Rightarrow f(y_1,\dots,y_n;\theta) = L(\theta)$ can be factored ∎

Back to the Poisson example...

$$L(\lambda) = \prod_{i=1}^{n} f(y_i|\lambda) = \prod_{i=1}^{n} \frac{e^{-\lambda}\lambda^{y_i}}{y_i!} = \frac{e^{-n\lambda}\lambda^{\Sigma y_i}}{\prod_{i=1}^{n} y_i!}$$

$$= \underbrace{\left[e^{-n\lambda}\cdot\lambda^{\Sigma y_i}\right]}_{g(u,\lambda)}\underbrace{\left[\frac{1}{\prod_{i=1}^{n} y_i!}\right]}_{h(y_1,\dots,y_n)}$$

$\boxed{u = \Sigma y_i}$ is a suff. stat. for λ by the factorization thm.

Back to normal dist. example...

$$L(\mu,\sigma^2) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\frac{1}{2\sigma^2}\Sigma(y_i-\mu)^2}$$

$$= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\frac{1}{2\sigma^2}\left[\Sigma(y_i-\bar{y})^2 + n(\bar{y}-\mu)^2\right]}$$

$$= \underbrace{\left[\frac{1}{(2\pi\sigma^2)^{n/2}} e^{-\frac{1}{2\sigma^2}(n-1)s^2}\cdot e^{-\frac{n}{2\sigma^2}(\bar{y}-\mu)^2}\right]}_{g(\bar{y},s^2,\mu,\sigma^2)}\cdot \underbrace{1}_{h(y_1\cdots y_n)}$$

$\Rightarrow \bar{y}, s^2$ are jointly suff. for μ, σ^2

e.g. 9.49 Let $Y_1 \cdots Y_n$ be a r.s. from $Y \sim unif(0,\theta)$

Find a sufficient stat. for θ

note: θ is a part of the support space for Y $(0<y<\theta)$ which means we need to use indicator fcts in $L(\theta)$

Let $I_{(0,\theta)}(y_i) = \begin{cases} 1, & \text{if } 0<y_i<\theta \\ 0, & \text{otherwise} \end{cases}$

$\Rightarrow f(y_i|\theta) = \frac{1}{\theta} \cdot I_{(0,\theta)}(y_i)$, $-\infty < y_i < \infty$

$\Rightarrow L(\theta) = \prod_{i=1}^{n} \frac{1}{\theta} \cdot I_{(0,\theta)}(y_i)$, $-\infty < y_i < \infty$

$= \frac{1}{\theta^n} \prod_{i=1}^{n} I_{(0,\theta)}(y_i)$, $-\infty < y_i < \infty$

note: $\prod_{i=1}^{n} I_{(0,\theta)}(y_i) = 1$ iff. $0 < y_{(1)}$ and $y_{(n)} < \theta$

$\Rightarrow \prod_{i=1}^{n} I_{(0,\theta)} = I_{(0,\infty)}(y_{(1)}) \cdot I_{(-\infty,\theta)}(y_{(n)})$

$$\Rightarrow L(\theta) = \underbrace{\frac{1}{\theta^n} I_{(-\infty,\theta)}(y_{(n)})}_{g(u,\theta)} \cdot \underbrace{I_{(0,\infty)}(y_{(1)})}_{h(y_1 \ldots y_n)}$$

$\Rightarrow$ $\boxed{u = Y_{(n)} \text{ is suff for } \theta}$

9.50 Let $Y_1, Y_2, \ldots, Y_n$ denote a r.s. from $Y \sim Unif(\theta_1, \theta_2)$
Find jointly sufficient statistics for θ_1, θ_2.

$f(y_i|\theta_1, \theta_2) = \frac{1}{\theta_2 - \theta_1}$, $\theta_1 < y_i < \theta_2$

or $= \frac{1}{\theta_2 - \theta_1} I_{(\theta_1, \theta_2)}(y_i)$, $-\infty < y_i < \infty$

$\Rightarrow L(\theta_1, \theta_2) = \prod_{i=1}^{n} f(y_i \mid \theta_1, \theta_2)$

$$= \frac{1}{(\theta_2 - \theta_1)^n} \underbrace{\prod_{i=1}^{n} I_{(\theta_1, \theta_2)}(y_i)}_{=1 \text{ if } y_{(1)} > \theta_1 \text{ and } y_{(n)} < \theta_2}$$

$$= \underbrace{\frac{1}{(\theta_2 - \theta_1)^n} I_{(\theta_1, \infty)}(y_{(1)}) \cdot I_{(-\infty, \theta_2)}(y_{(n)})}_{g(y_{(1)}, y_{(n)}, \theta_1, \theta_2)} \cdot \underbrace{1}_{h(y_1 \cdots y_n)}$$

$\Rightarrow$ by Factorization Thm, $Y_{(1)}$ and $Y_{(n)}$ are jointly sufficient for θ_1 and θ_2

Exponential Class

Def A density fct. is said to be a member of the regular exponential class if it can be expressed in the form

$$f(y \mid \theta) = c(\theta) h(y) e^{-g(\theta) \cdot d(y)}, \quad a \le y \le b$$

where a and b do not depend on θ

9.45 $f(y; \theta) = c(\theta) \cdot h(y) e^{-g(\theta) \cdot d(y)}, \quad a \le y \le b$

Show $\sum_{i=1}^{n} d(y_i)$ is sufficient for θ

$$L(\theta) = \prod_{i=1}^{n} c(\theta) \cdot h(y_i) e^{-g(\theta) \cdot d(y_i)}$$

$$= c(\theta)^n \cdot \left[\prod_{i=1}^{n} h(y_i)\right] e^{-g(\theta)\sum_{i=1}^{n} d(y_i)}$$

$$= \underbrace{\prod_{i=1}^{n} h(y_i)}_{h(y_1 \cdots y_n)} \underbrace{\left[c(\theta)^n \cdot e^{-g(\theta)\sum d(y_i)}\right]}_{g(u,\theta)}$$

$\Rightarrow \boxed{u = \sum_{i=1}^{n} d(y_i)}$ is sufficient for θ by the factorization Thm

Back to Poisson dist ...

Let $Y_1 \cdots Y_n$ be a r.s. from $Y \sim POI(\lambda)$

a) Show the Poisson is regular exponential class

$$p(y_i|\lambda) = \frac{e^{-\lambda}\lambda^{y_i}}{y_i!} = \boxed{e^{-\lambda} \cdot \left(\frac{1}{y_i!}\right) e^{-y_i \cdot \ln\lambda}, \quad y = 0,1,2,\ldots}$$

$c(\lambda) = e^{-\lambda}$ $h(y_i) = \frac{1}{y_i!}$ $g(\lambda) = \ln\lambda$ $d(y_i) = y_i$ and support does not depend on λ

b) Find a suff. stat. for λ

By result of 9.45 $\Rightarrow u = \sum_{i=1}^{n} d(y_i) = \boxed{\sum_{i=1}^{n} y_i}$ is suff. for λ