Let $T_1, \dots, T_k$ be random variables with means $\theta_1, \dots, \theta_k$ and define $\underline{T} = (T_1, \dots, T_k)$ and $\underline{\theta} = (\theta_1, \dots, \theta_k)$

Assume $g(\underline{t})$ is differentiable and let $g_i'(\underline{\theta}) = \frac{\partial}{\partial \theta_i} g(\underline{t}) \Big|_{\underline{\theta} = (\theta_1, \dots, \theta_k)}$

Then $g(\underline{t}) = g(\underline{\theta}) + \sum_{i=1}^{k} g_i'(\underline{\theta})(T_i - \theta_i) + R$

and $E[g(\underline{T})] \approx g(\underline{\theta}) + \sum_{i=1}^{k} g_i'(\underline{\theta}) E[T_i - \theta_i] = g(\underline{\theta})$

$$Var[g(\underline{T})] \approx Var\left[\left(\sum_{i=1}^{k} g_i'(\underline{\theta})(T_i - \theta_i)\right)\right] = \sum_{i=1}^{k} [g_i'(\underline{\theta})]^2 Var[T_i]$$

$$+ 2 \sum_{i \neq j} g_i'(\underline{\theta}) g_j'(\underline{\theta}) Cov(T_i, T_j)$$

Theorem 5.5.24. The Delta method

$\{Y_n\}_{n=1}^{\infty}$ satisfies $\sqrt{n}(Y_n - \theta) \xrightarrow{D} N(0, \sigma^2)$ $\left(\Leftrightarrow \frac{Y_n - \theta}{\frac{\sigma}{\sqrt{n}}} \xrightarrow{D} N(0,1) \right.$

Assume $g'(\theta)$ exists and is different from 0. Then

$\sqrt{n}[g(Y_n) - g(\theta)] \xrightarrow{D} N(0, \sigma^2 (g'(\theta))^2]$

Proof. A first order Taylor expansion gives

$$g(Y_n) = g(\theta) + g'(\theta)(Y_n - \theta) + \underbrace{R(Y_n, \theta)}_{= C_1 (Y_n - \theta)^2 + C_2 (Y_n - \theta)^3 + \dots}$$

$$\sqrt{n}[g(Y_n) - g(\theta)] = g'(\theta)\sqrt{n}(Y_n - \theta) + \sqrt{n}(Y_n - \theta)(C_1(Y_n - \theta) + C_2(Y_n - \theta)^2 + \dots$$

What happens to $Y_n - \theta$ as $n \to \infty$

$X_n = \frac{\sqrt{n}}{\sigma}(Y_n - \theta) \xrightarrow{D} N(0,1) \Rightarrow \underbrace{\frac{\sigma}{\sqrt{n}}}_{a_n} X_n = Y_n - \theta \xrightarrow{D} (\lim_{n \to \infty} a_n) N(0,1) = 0$

i.e. $Y_n - \theta \xrightarrow{D} 0$ and $Y - \theta \xrightarrow{P} 0. \Rightarrow \sqrt{n} R(Y_n, \theta) \xrightarrow{P} 0$

__Example__ $X_1, \dots, X_m$ i.i.d. $E[X_i] = \mu$, $E[\bar{X}_m] = \mu$.

$Var[X_1] = \sigma^2$

Let $g(X) = \frac{1}{X}$, $g(\bar{X}_m) = \frac{1}{\bar{X}_m}$

Then

$$\sqrt{m}\left(\frac{1}{\bar{X}_m} - \frac{1}{\mu}\right) \xrightarrow{D} N\left(0, \frac{1}{\mu^4} Var(X_i)\right)$$

Assume $Var[X_i]$ is unknown

$$Var\left[\frac{1}{\bar{X}}\right] \approx \left(\frac{1}{\bar{X}}\right)^4 S_m^2$$

Then
$$\frac{\sqrt{m}\left(\frac{1}{\bar{X}_m} - \frac{1}{\mu}\right)}{\frac{1}{\bar{X}^2} S_m} = \underbrace{\frac{\sqrt{m}\left(\frac{1}{\bar{X}_m} - \frac{1}{\mu}\right)}{\frac{\sigma}{\mu^2}}}_{\xrightarrow{D} N(0,1)} \cdot \underbrace{\frac{\sigma}{S_m} \cdot \frac{\bar{X}_m^2}{\mu^2}}_{\xrightarrow{P} 1} \xrightarrow{D} N(0,1).$$

__Theorem 5.5.26__ Second order Delta method.

$\{Y_m\}_{m=1}^{\infty}$ satisfies $\sqrt{m}(Y_m - \theta) \xrightarrow{D} N(0, \sigma^2)$. Suppose for given θ

$g'(\theta) = 0$ and $g''(\theta)$ exists $\neq 0$. Then

$$m[g(Y_m) - g(\theta)] \xrightarrow{D} \frac{\sigma^2}{2} g''(\theta) \chi^2(1)$$

__Proof__. A second order Taylor expansion gives.

$$g(Y_m) - g(\theta) = \frac{g''(\theta)}{2}(Y_m - \theta)^2 + \left(c_1 (Y_m - \theta)^3 + c_2 (Y_m - \theta)^4 + \dots\right)$$

$$\Rightarrow m[g(Y_m) - g(\theta)] = \frac{g''(\theta)}{2} \cdot \sigma^2 \underbrace{\left(\frac{m(Y_m - \theta)^2}{\sigma^2}\right)}_{\xrightarrow{D} \chi^2(1)} + \underbrace{m \frac{\sigma^2}{\sigma^2}(Y_m - \theta)^2}_{\xrightarrow{D} \sigma^2 \chi^2(1)} \underbrace{\left(c_1 (Y_m - \theta) + c_2 (Y_m - \theta)^2 + \dots\right)}_{\xrightarrow{P} 0}.$$

Thereby $m[g(Y_m) - g(\theta)] \xrightarrow{D} \frac{g''(\theta)}{2} \sigma^2 \chi^2(1)$

Chapter 6. Principles of Data Reduction

An example

X_1, X_2 Bernoulli and independent

i.e $P(X_1=x) = P(X_2=x) = \begin{cases} 1-\theta, & x=0 \\ \theta, & x=1 \end{cases}$ $\theta \in (0,1)$

$T = X_1 + X_2$

Experimenter 2 get the information $T = 0, 1$ or 2

X-space Experimenter 1	T-space Experimenter 2
$(0,0)$	0
$(1,0)$	1
$(1,0)$	1
$(1,1)$	2
$(0,1)$	1

$$P((X_1, X_2) = (0,0) \mid T=0) = 1$$

$$P((X_1, X_2) = (1,0) \mid T=1) = \frac{P((X_1,X_2)=(1,0) \cap T=1)}{P(T=1)} = \frac{\theta(1-\theta)}{2\theta(1-\theta)} = \frac{1}{2}$$

$$= P((X_1,X_2)=(0,1) \mid T=1)$$

$$P((X_1,X_2) = (1,1) \mid T=2) = 1$$

Let us generate (X_1°, X_2°) as $\begin{cases} (0,0) & \text{if } T=0 \\ (1,0) \text{ or } (0,1) \text{ each with probabilities } \tfrac{1}{2} & \text{if } T=1 \\ (1,1) & \text{if } T=2 \end{cases}$

Note. No information about θ is needed to generate (X_1°, X_2°)

We have

$$P((X_1^\circ, X_2^\circ) = (0,0)) = P((X_1^\circ, X_2^\circ) = (0,0) \cap T(X_1, X_2) = 0)$$

$$= P((X_1^\circ, X_2^\circ) = (0,0) \mid T(X_1, X_2) = 0) \cdot P(T(X_1, X_2) = 0) = (1-\theta)^2 =$$

$$P((X_1, X_2) = (0,0)).$$

Similarly for $(X_1^\circ, X_2^\circ) = (1,0), (0,1)$ and $(1,1)$

Hence (X_1, X_2) and (X_1°, X_2°) have the same distribution, but are normally not identical, but inference about θ can be equally well made from (X_1°, X_2°)

<u>Definition 6.2.1</u>

$T(\underline{X})$ is a sufficient for θ if the conditional distribution of $\underline{X}$ given $T(\underline{X})$ does not depend on θ.

A sufficient statistics for a parameter θ is a statistic that in a certain sense capture all the information about θ in a sample.

<u>Discrete case</u>

In general experimenter 2 who knows $T(\underline{X}) = t$ can generate a distribution $P(\underline{X} = \underline{y} \mid T(\underline{X}) = t)$ on the set $A_{T(\underline{X})} = \{\underline{y} : T(\underline{y}) = t\}$

A random variable $\underline{Y}$ with this distribution satisfies

$$P(\underline{Y} = \underline{y} \mid T(\underline{X}) = t) = P(\underline{X} = \underline{y} \mid T(\underline{X}) = t)$$