dist	$I_1(\theta)$	AVAR $Var(\hat\theta)$	
Weibull(θ,τ) τ known	$\frac{\tau^2}{\theta^2}$	$\frac{\theta^2}{n\tau^2}$	
$U(0,\theta)$ not a 1PREF	$\frac{1}{\theta^2}$	$\frac{n\theta^2}{(n+1)^2(n+2)}$	$\approx \frac{\theta^2}{n^2} \neq \frac{1}{nI_1(\theta)}$ not asymptotically normal
$N(\mu,\sigma^2)$	$I_1(\mu)=\frac{1}{\sigma^2}$	$\frac{\sigma^2}{n}$	(ACOV = cov in this case) $cov(\hat\mu,\hat\sigma^2)=0$
	$I_1(\sigma^2)=\frac{1}{2\sigma^4}$	$\frac{2\sigma^4}{n}$	
pois(λ)	$\frac{1}{\lambda}$	$\frac{\lambda}{n}$	
negative binomial(r,β) r known	$\frac{r}{\beta(1+\beta)}$	$\frac{\beta(1+\beta)}{nr}$	
bin(q,m) m known	$\frac{m}{q(1-q)}$	$\frac{q(1-q)}{nm}$	
Gamma(α,θ) α known	$\frac{\alpha}{\theta^2}$	$\frac{\theta^2}{n\alpha}$	see E2 rev 73)
Geom(β)	$\frac{1}{\beta(1+\beta)}$	$\frac{\beta(1+\beta)}{n}$	

63] Often $X_1,\ldots,X_n$ are iid and $\hat\theta = c\bar{X}$ $\rightarrow$ $I_1(\theta)=\frac{1}{c^2V(X)}$

$$\sqrt{n}\,(\bar{X}-E(X)) \xrightarrow{D} N[0, V(X)]$$

$$\sqrt{n}\,(c\bar{X}-cE(X)) = \sqrt{n}(\hat\theta-\theta) \xrightarrow{D} N\left(0,\frac{1}{I_1(\theta)}\right) = N[0, c^2V(X)]$$

ex] $X\sim Pois(\lambda)$, $X\sim EXP(\theta)$, $X\sim N(\mu,\sigma^2)$ with $\theta=\mu$ and

$X \sim \text{Geom}(\beta)$ have $c = 1$

$X \sim \text{bin}(q, m)$ has $\hat{q} = \frac{\bar{X}}{m}$ so $c = \frac{1}{m}$
m known

$X \sim G(\alpha, \theta)$ has $\hat{\theta} = \frac{\bar{X}}{\alpha}$ so $c = \frac{1}{\alpha}$.
α known

$X \sim$ negative binomial (r, β) has $\hat{\beta} = \frac{\bar{X}}{r}$ so $c = \frac{1}{r}$.
r known

ex] $X \sim U(0, \theta)$ has $f(x) = \frac{1}{\theta}$ $\quad 0 < x < \theta$

not a 1PREF

$$\ln f(x) = -\ln(\theta)$$

$$\frac{d}{d\theta} \ln f(x) = \frac{-1}{\theta}$$

$$E\left(\left[\frac{d}{d\theta} \ln f(x)\right]^2\right) = E\left(\frac{1}{\theta^2}\right) = \frac{1}{\theta^2} = I_1(\theta).$$

64] Let $I_1(\underline{\theta})$ be the information matrix when $n = 1$. So $I_{ij} =$

$$E\left[\frac{\partial}{\partial\theta_i} \ln f(x|\underline{\theta}) \frac{\partial}{\partial\theta_j} \ln f(x|\underline{\theta})\right] = $$

under regularity conditions such as KP-REF

$$-E\left[\frac{\partial^2}{\partial\theta_i \partial\theta_j} \ln f(x|\underline{\theta})\right].$$ Under regularity conditions $\sqrt{n}(\hat{\underline{\theta}} - \underline{\theta}) \xrightarrow{D} N_k(\underline{0}, I_1^{-1}(\underline{\theta}))$

and the asymptotic covariance

matrix of $\hat{\theta}$ is

$$\underbrace{\mathrm{cov}(\hat{\theta})}_{\mathrm{Acov}(\hat{\theta})} = I_n^{-1}(\underline{\theta}) = [n I_1(\underline{\theta})]^{-1}.$$

65] <u>Delta method:</u> if

$\sqrt{n}(T_n - \theta) \xrightarrow{D} N(0, \sigma^2)$ and $g'(\theta) \neq 0$,

then $\sqrt{n}\left(g(T_n) - g(\theta)\right) \xrightarrow{D} N(0, \sigma^2 [g'(\theta)]^2)$.

Apply to CLT with $T_n = \bar{X}$ and $\theta = \mu$

or to the MLE limit theorem with

$T_n = \hat{\theta}_n$ and $g(\theta) = \tau(\theta)$.

66] <u>Delta method:</u> If $g'(\theta) \neq 0$ and

$\mathrm{var}(\hat{\theta})$ is the asymptotic variance of $\hat{\theta}$,

then the asymptotic variance of $g(\hat{\theta})$

is $[g'(\hat{\theta})]^2 \, \mathrm{var}(\hat{\theta})$.

67] Delta method: Let $\underline{\hat{\theta}} = (\hat{\theta}_1, \hat{\theta}_2)$ have

asymptotic covariance matrix

$$\mathrm{cov}(\underline{\hat{\theta}}) = \begin{pmatrix} \mathrm{var}(\hat{\theta}_1) & \mathrm{cov}(\hat{\theta}_1, \hat{\theta}_2) \\ \mathrm{cov}(\hat{\theta}_1, \hat{\theta}_2) & \mathrm{var}(\hat{\theta}_2) \end{pmatrix} \underset{\text{if } \underline{\hat{\theta}} \text{ is the MLE}}{=} I_n^{-1}(\underline{\theta})$$

Let $g : \mathbb{R}^2 \to \mathbb{R}$ have nonzero partial derivatives. Then the asymptotic variance of $g(\hat{\theta}_1, \hat{\theta}_2)$ is

$$\operatorname{var}\left[g(\hat{\theta}_1, \hat{\theta}_2)\right] = \left(\frac{\partial g(\underline{\theta})}{\partial \theta_1}\right)^2 \overbrace{\operatorname{var}(\hat{\theta}_1)}^{\operatorname{cov}(\hat{\theta}_1, \hat{\theta}_1)}$$

$$+ 2\left(\frac{\partial g(\underline{\theta})}{\partial \theta_1}\right)\left(\frac{\partial g(\underline{\theta})}{\partial \theta_2}\right) \operatorname{cov}(\hat{\theta}_1, \hat{\theta}_2) +$$

$$\left(\frac{\partial g(\underline{\theta})}{\partial \theta_2}\right)^2 \operatorname{var}(\hat{\theta}_2). =$$

$LN(\mu, \sigma^2)$
$N(\mu, \sigma^2)$
sometimes
$\hat{\sigma}_{12} = 0$

$$\begin{pmatrix} \frac{\partial g(\underline{\theta})}{\partial \theta_1} & \frac{\partial g(\underline{\theta})}{\partial \theta_2} \end{pmatrix} \begin{pmatrix} \hat{\sigma}_{11} & \hat{\sigma}_{12} \\ \hat{\sigma}_{21} & \hat{\sigma}_{22} \end{pmatrix} \begin{pmatrix} \frac{\partial g(\underline{\theta})}{\partial \theta_1} \\ \frac{\partial g(\underline{\theta})}{\partial \theta_2} \end{pmatrix}$$

$$= \left[\frac{\partial g(\underline{\theta})}{\partial \theta_1}\hat{\sigma}_{11} + \frac{\partial g(\underline{\theta})}{\partial \theta_2}\hat{\sigma}_{21} \quad \frac{\partial g(\underline{\theta})}{\partial \theta_1}\hat{\sigma}_{12} + \frac{\partial g(\underline{\theta})}{\partial \theta_2}\hat{\sigma}_{22}\right] \begin{pmatrix} \frac{\partial g(\underline{\theta})}{\partial \theta_1} \\ \frac{\partial g(\underline{\theta})}{\partial \theta_2} \end{pmatrix},$$

<u>Mnemonic:</u> $\operatorname{cov}(\hat{\theta}_i, \hat{\theta}_j)$ has coefficient

$\dfrac{\partial g(\underline{\theta})}{\partial \theta_i} \dfrac{\partial g(\underline{\theta})}{\partial \theta_j}$ where $\operatorname{cov}(\hat{\theta}_i, \hat{\theta}_i) = \operatorname{var}(\hat{\theta}_i)$

and $\operatorname{cov}(\hat{\theta}_1, \hat{\theta}_2) = \operatorname{cov}(\hat{\theta}_2, \hat{\theta}_1)$ appears twice.

ex] $X \sim$ single parameter Pareto

$(\theta = 100, \alpha)$, $\hat{\alpha} = 3$,

$\text{var}(\hat{\alpha}) = 0.5$. Determine the

variance of the estimate of $P(X < 200)$.

AVAR

x contin

Soln $P(X < 200) = F(200) =$

$$1 - \left(\frac{100}{200}\right)^{\alpha} = 1 - \left(\frac{1}{2}\right)^{\alpha} = g(\alpha).$$

The MLE of $g(\alpha)$ is $g(\hat{\alpha})$ by invariance

$\frac{d}{dx} a^x = a^x \ln(a)$

and $\overbrace{\text{Var}(g(\hat{\alpha}))}^{\text{AVAR}} = [g'(\alpha)]^2 \, \text{Var}(\hat{\alpha}) =$

$$= \left(-\left(\tfrac{1}{2}\right)^{\alpha} \ln\left(\tfrac{1}{2}\right)\right)^2 0.5 \underset{\hat{\alpha}}{\approx} \left[\left(\tfrac{1}{2}\right)^3 \ln\left(\tfrac{1}{2}\right)\right]^2 0.5$$

$= \boxed{0.003754}$.

68] Know: when computing $\text{Var}(g(\hat{\theta}))$,

plug in $\hat{\theta}$ to get a real number.

when computing $\text{cov}(\hat{\underline{\theta}})$, plug in

$\hat{\underline{\theta}}$ to get real entries in the matrix.

$\sigma_{12}=\sigma_{21}$ symmetric

69} If $\sqrt{n}\left(\begin{pmatrix}\hat{\theta}_1\\ \hat{\theta}_2\end{pmatrix}-\begin{pmatrix}\theta_1\\ \theta_2\end{pmatrix}\right)\xrightarrow{D} N_2\left[\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}\sigma_{11} & \sigma_{12}\\ \sigma_{21} & \sigma_{22}\end{pmatrix}\right]$,

then $\sqrt{n}\,(\hat{\theta}_i-\theta_i)\xrightarrow{D} N(0,\sigma_{ii})$ for $i=1,2$.

ex} $X\sim$ Pareto(α,θ), $(\hat{\alpha},\hat{\theta})=(3,100)$,

covariance matrix $=\begin{pmatrix}0.6 & 0.2\\ 0.2 & 0.4\end{pmatrix}$ $\left(=\text{cov}(\hat{\alpha},\hat{\theta})\right)$.

Find the variance of the estimate of $P(X<200)$.

Soln: Let $g(\alpha,\theta)=P(X<200)=F(200)$

$=1-\left(\frac{\theta}{\theta+200}\right)^{\alpha}$.

$\frac{d}{dx}a^x=a^x\ln a$

$\frac{\partial g}{\partial\alpha}=-\left(\frac{\theta}{\theta+200}\right)^{\alpha}\ln\left(\frac{\theta}{\theta+200}\right)=-\left(\frac{1}{3}\right)^3\ln\left(\frac{1}{3}\right)=0.04069$

$$\frac{\partial g}{\partial \theta} = -\alpha\left(\frac{\theta}{\theta+200}\right)^{\alpha-1}\frac{(\theta+200)1-\theta(1)}{(\theta+200)^2}$$

Quotient rule $\dfrac{dn' - nd'}{d^2}$

$$= -\alpha\left(\frac{\theta}{\theta+200}\right)^{\alpha-1}\frac{200}{(\theta+200)^2} \approx$$

$$-3\left(\frac{1}{3}\right)^2\frac{200}{(300)^2} = -0.0007407.$$

So $Var\left(P(X<200)\right) = Var\left(g(\hat{\alpha},\hat{\theta})\right)$

"easy" if both #'s are given

$$= (.04069)^2\, 0.6 - 2\,(.04069)(.0007407)\, 0.2$$

$$+ (.0007407)^2\, 0.4 = \boxed{0.0009816}$$

Keep at least 4 digits past leading 0's. *

70) If $\sqrt{n}\,(\hat{\theta}-\theta) \xrightarrow{D} N(0, \sigma^2)$,

then $Var(\hat{\theta}) = \frac{\sigma^2}{n}$ and a

large sample $100(1-\frac{p}{2})\%$ confidence interval

CI for θ is $\hat{\theta} \pm z_p \sqrt{\widehat{Var}(\hat{\theta})} =$

$\left(\hat{\theta} - z_p \frac{\hat{\sigma}}{\sqrt{n}},\ \hat{\theta} + z_p \frac{\hat{\sigma}}{\sqrt{n}}\right)$ where

$P(Z > z_p) = \frac{p}{2}$ or $P(Z \le z_p) = 1 - \frac{p}{2}$ if $Z \sim N(0,1)$

(earlier had $P(Z \le z_p) = p$).

71)

CI	90%	95%	99%
z_p	1.645	1.96	2.576

72) For the MLE, typically $\sigma^2 = \frac{1}{I_1(\theta)}$ and

$$var(\hat{\theta}) \doteq \frac{\sigma^2}{n} = \frac{1}{I_n(\theta)} = \frac{1}{n I_1(\theta)}. \text{ So}$$

$$\frac{\hat{\sigma}}{\sqrt{n}} = \frac{1}{\sqrt{n I_1(\hat{\theta})}} = \frac{1}{\sqrt{I_n(\hat{\theta})}}. \text{ So}$$

the $100(1-\frac{p}{2})\%$ CI for θ is

$$\hat{\theta} \pm z_p \frac{1}{\sqrt{n I_1(\hat{\theta})}} = \hat{\theta} \pm z_p \frac{1}{\sqrt{I_n(\hat{\theta})}}.$$

Since $\sqrt{n}\left[\tau(\hat{\theta}) - \tau(\theta)\right] \xrightarrow{D} N\left(0, \frac{(\tau'(\theta))^2}{I_1(\theta)}\right)$,

the $100(1-\frac{p}{2})\%$ CI for $\tau(\theta)$

is $\tau(\hat{\theta}) \pm z_p \sqrt{\frac{[\tau'(\hat{\theta})]^2}{n I_1(\hat{\theta})}} = \tau(\hat{\theta}) \pm z_p \sqrt{\frac{[\tau'(\hat{\theta})]^2}{I_n(\hat{\theta})}}$

The $100(1-\frac{\rho}{2})\%$ CI for

$g(\theta_1,\theta_2)$ is $g(\hat{\theta}_1,\hat{\theta}_2) \pm z_\rho \sqrt{\widehat{Var}(g(\hat{\theta}_1,\hat{\theta}_2))}$

plug $\hat{\theta}_1, \hat{\theta}_2$ into $Var(g(\hat{\theta}_1,\hat{\theta}_2))$.

73] *Let m = number of uncensored observations,
c = number of censored observations,
$n = m+c$, d_i is the truncation point for each observation (0 if untruncated))
x_i is the observation if uncensored
or the censoring point (u_i) if censored.
The following formulas work if left truncation and right censoring are present or not.

a) $Exp(\theta)$: $\hat{\theta} = \dfrac{\sum_{i=1}^{n}(x_i - d_i)}{m} = \dfrac{\sum_{i=1}^{n} y_i^{\rho}}{m}$

b) Weibull fixed τ: $\hat{\theta} = \left(\dfrac{\sum_{i=1}^{n}(x_i^{\tau} - d_i^{\tau})}{m}\right)^{\frac{1}{\tau}}$
|| known

c) pareto fixed θ: $\hat{\alpha} = \dfrac{-m}{\sum_{i=1}^{n} \ln\left(\frac{\theta+d_i}{\theta+x_i}\right)}$

d) single parameter pareto, fixed θ:

$$\hat{\alpha} = \frac{-m}{\sum_{i=1}^{n} \ln\left(\frac{\max(\theta, d_i)}{x_i}\right)}$$

See Ez rev 78)

54] Suppose the $U(0,\theta)$ dist is used for grouped or censored data. Let $[c, f)$ be the highest interval with count $n_c > 0$. ($f = \infty$ is allowed, c could be the censoring value). Let $m = \#$ obs's $< c$ so $n = m + n_c$.

Then $\hat{\theta} = \min\left(\frac{n}{m}c, f\right)$. Note that $\hat{\theta}$ is found by matching $P(X<c) = p = \frac{c}{\theta} \overset{set}{=} \hat{p} = \frac{m}{n}$.

$(0, c)$ m } censored
$[c, \infty)$ $n-m$ } data ,

(c_0, c_1) n_1
(c_1, c_2) n_2
$\vdots$
(c_{k-1}, c_k) $n_k > 0$

or

(c_0, c_1) n_1
(c_1, c_2) n_2
$\vdots$
(c_{k-1}, c_k) $n_k > 0$
(c_k, ∞) 0

$n = \sum_{i=1}^{k} n_i$, $m = n - n_k$
$c = c_{k-1}$

ex] Insurance has a deductible 500 and maximum covered loss 10000. Reported