$$= \left(1 - e^{\frac{-1000}{\theta}}\right)^{15} \left(e^{\frac{-1000}{\theta}}\right)^{11} = (1-x)^{15} x^{11} = h(x)$$ M404 20

with $x = e^{\frac{-1000}{\theta}}$,

$\ln L(\theta) = 15 \ln(1-x) + 11 \ln(x) = g(x)$

$\frac{d}{dx} g(x) = 15 \frac{1}{1-x}(-1) + \frac{11}{x} \stackrel{set}{=} 0$ or

maximizing $g(x)$ will maximize $\ln L(\theta)$

$-15x + 11(1-x) = 0$

$-15x + 11 - 11x = 0$

$26x = 11$

$x = \frac{11}{26} = e^{\frac{-1000}{\theta}}$

$-\frac{1000}{\theta} = \ln \frac{11}{26}$

$$\hat{\theta} = \frac{-1000}{\ln\left(\frac{11}{26}\right)} = \boxed{1162.5786}$$

Note that even simple problems are rather hard to solve.

36) Suppose $x_1, \ldots, x_n$ are left truncated at d (observation x_i is only observed if $x_i > d$). Then $L(\underline{\theta} | x_i) = \frac{f(x_i)}{1-F(d)}$. If the data are right truncated (obs. x_i only observed if $x_i < U$),

then $L(\underline{\theta} \mid x_i) = \dfrac{f(x_i)}{F(U)}$.

37) a) If x_i is censored above at U,

$$L(\underline{\theta} \mid x_i) = 1 - F(U).$$

b) If x_i is censored below at d,

$$L(\underline{\theta} \mid x_i) = F(d)$$

38)

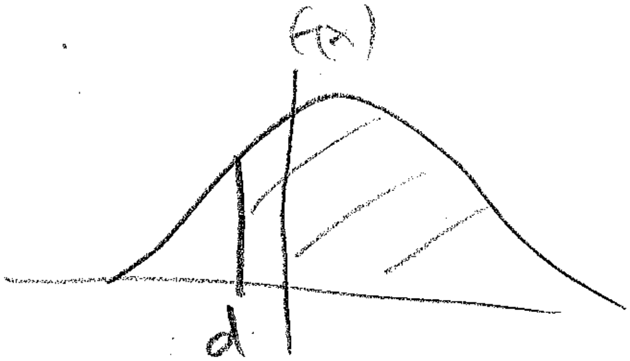
	x_i	$L(\underline{\theta} \mid x_i)$	
i)	$X_i \sim$ discrete X	$p(x_i)$	M48%, 580
ii)	$X_i \sim$ contin X	$f(x_i)$	
iii)	X_i censored above at U	$\begin{cases} 1-F(U) & \text{censored } x_i \text{ recorded as } U \\ f(x_i) & x_i \le U \end{cases}$	
iv)	x_i censored below at d	$\begin{cases} F(d) & \text{censored } x_i \text{ recorded as } d \\ f(x_i) & x_i > d \end{cases}$	rare
v)	x_i from X truncated above at U	$\dfrac{f(x_i)}{F(U)}$	rare
vi)	x_i from X truncated below at d	$\dfrac{f(x_i)}{1-F(d)}$	
vii)	$c_{j-1} < x_i \le c_j$ grouped data	$F(c_j) - F(c_{j-1})$	contin X_i have useful F

Note that $S(x) = 1 - F(x)$.

39) Some of these concepts can be combined.

Note! my webpage 2 course for Statistical Theory Ch 10 does $L(\theta)$ $\log L(\theta)$ $\frac{d}{d\theta} \log L(\theta)$ and solves for the MLE $\hat{\theta}$ for about 23 distributions.

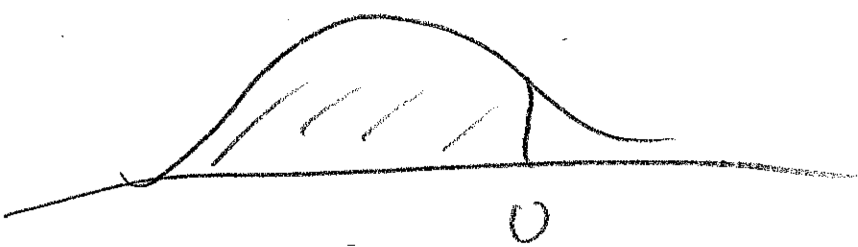
See M404 A 2D



$$\int_d^\infty f(x)\,dx = S(d) \neq 1$$

So if W is from X left truncated at d,

then $f_W(w) = \dfrac{f(x)}{S(d)} = \begin{cases} \dfrac{f(x)}{1-F(d)} & , x > d \\ 0 & , x < d \end{cases}$



$$\int_{-\infty}^{U} f(x) = F(U) \neq 1$$

So if W is from X right truncated at U, then

$f_W(w) = \begin{cases} \dfrac{f(x)}{F(U)} & x \leq U \\ 0 & x > U \end{cases}$

a) If data $x_1, \ldots, x_m, \underbrace{u, \ldots, u}_{n-m}$ are left truncated and right censored,

$L(\theta|x_i) = \dfrac{f(x_i)}{1-F(d)}$, $\prod_{i=1}^{n-m} L(\theta|u) = \left(\dfrac{1-F(u)}{1-F(d)}\right)^{n-m}$ and

$$L(\theta) = \left[\prod_{i=1}^{m} \frac{f(x_i)}{[1-F(d)]}\right] \left[\frac{1-F(u)}{1-F(d)}\right]^{n-m} = \frac{\left[\prod_{i=1}^{m} f(x_i)\right] [1-F(u)]^{n-m}}{[1-F(d)]^{n}}$$

b) If n_j grouped data are between d and c_j in presence of truncation at d, then these n_j terms contribute

$$\left[\frac{F(c_j) - F(d)}{1-F(d)}\right]^{n_j} \text{ to } L(\theta).$$

40] If there is a deductible d and a policy limit = maximum payment of $u-d$, then the maximum covered loss $u = u-d+d$.

For a coinsurance policy, the insurance company pays αX of the loss for some $\alpha \in (0,1]$. For coinsurance with a deductible, the insurance company pays $\alpha(X-d)_+$.

41] Using the above notation, if Y^P is the per payment RV, and X is the loss RV, then $Y^P = \alpha(X-d)$.

So $X = \frac{Y^P}{\alpha} + d$. $\to$ {only $X>d$ are observed}

Here X is left truncated (at d) and right censored (at u). Note that $d=0$, $\alpha=1$, and $u=\infty$ are possible (no deductible, no coinsurance, no maximum payment).
Note that Y^P is right censored at $\alpha(u-d)$.

ex] An insurance coverage had deductible 1000, maximum covered loss 50000 and 90% coinsurance. There are 20 payments (ins. company) $x_1, \ldots, x_{20}$ of amounts less than 44100 and 5 payments of 44100. Write the likelihood function for the loss distribution. Assume the parameter is θ.

soln] The x_i are for Y^P. Want $L(\theta)$ for X. So $X = \frac{Y^P}{0.9} + 1000$ has $w_i = \frac{x_i}{0.9} + 1000$.

Note that $0.9\,(50000-1000) = 44100$ so these observations were censored payments Y^P. For X, the censored values are 50000, and X is left truncated at 1000

$$L(\theta) = \frac{\left(\prod_{i=1}^{20} f\left(\frac{x_i}{0.9}+1000\right)\right)\left[1-F(50000)\right]^5}{\left[S(1000)\right]^{25}}$$

$\leftarrow$ 5 censored, 20 not censored

$\leftarrow$ all 25 obs's are truncated

$$= \frac{\left[\prod_{i=1}^{20} f\left(\frac{x_i}{0.9}+1000\right)\right]\left[S(50000)\right]^5}{\left[S(1000)\right]^{25}}$$

know that $S(x) = 1 - F(x)$.

42] For individual discrete data (iid)

$$L(\theta) = \prod_{i=1}^{n} p(x_i) = \prod_{j=1}^{m} [p(x_j)]^{n_j}$$

where $n_j = \#$ times $x_i = x_j$ for $j = 1, \ldots, m$

ex)

x_j	number
0	41
1	35
2	17
3	5
4	2

out ← rpois(100, 1)

sum(out == 0)

sum(out == 4)

$\Sigma x_i = 0 + 35 + 2(17) + 3(5) + 4(2) = 92$

If these 100 obs's are iid from a Pois(λ) dist,

$$\text{then } L(\lambda) = (e^{-\lambda})^{41} (e^{-\lambda}\lambda)^{35} \left(\frac{e^{-\lambda}\lambda^2}{2}\right)^{17} \left(\frac{e^{-\lambda}\lambda^3}{6}\right)^5 \left(\frac{e^{-\lambda}\lambda^4}{24}\right)^2$$

$$\propto c\, e^{-\lambda 100} \lambda^{92}$$

$$\ln L(\lambda) = d - \lambda 100 + 92 \ln(\lambda)$$

$$\frac{d \ln L(\lambda)}{d\lambda} = -100 + \frac{92}{\lambda} \overset{set}{=} 0 \text{ or } 100\lambda = 92, \ \underbrace{\hat{\lambda} = \frac{92}{100} = \bar{X}}_{\text{E1 rev}}$$

43] For independent data $x_1, \dots, x_n$,

$$L(\underline{\theta}) = \prod_{i=1}^{n} f_{X_i}(x_i) \quad \text{or} \quad L(\underline{\theta}) = \prod_{i=1}^{n} P_{X_i}(x_i),$$

the X_i need not have the same distr.

Then $\ell(\underline{\theta}) = \ln L(\underline{\theta}) = \sum_{i=1}^{n} \ln[f_{X_i}(x_i)]$ or $\sum_{i=1}^{n} \ln[P_{X_i}(x_i)]$.

44] p343 * When finding $L(\underline{\theta})$, a) for the numerator, use $f(x_i)$ if x_i is known. If it is only known that x_i is between y and z, use $F(z) - F(y)$ (grouped data or (right) data censored at u with $y = u$, $z = \infty$ so $F(\infty) - F(u) = 1 - F(u) = S(u)$).

b) For the denominator, let d be the truncation point (use $d = 0$ if there is no (left) truncation). Then the denominator is $1 - F(d) = S(d)$.

45] Could have data x_i uncensored, or y_i censored, u_i where the censoring value $y_i = u_i$ and truncation value ≑ deductible d_i vary with i.

ex p344 obs	x_i or y_i (w_i)	d_i	$L(\theta \mid w_i)$	
1	$y=0.1$	0	$1-F(0.1)$	censored at 0.1
⋮				
4	$x=0.8$	0	$f(0.8)$	uncensored
⋮				
31	$y=5$	0.3	$\frac{1-F(5)}{1-F(0.3)}$	censored and truncated
⋮				
33	$x=4.1$	1	$\frac{f(4.1)}{1-F(1)}$	uncensored but truncated

end 403 review

§15.3 46) p414, 609 Let $f(x)$ be a pdf or pmf. Assume the support of X does not depend on any unknown parameters. The family of pdfs or pmfs is a k parameter exponential family if

$$f(x|\underline{\theta}) = h(x)\, c(\underline{\theta}) \exp\left[\sum_{i=1}^{k} w_i(\underline{\theta})\, t_i(x)\right] \quad (*)$$

where k is the smallest integer where (*) holds.

47) For a one parameter exponential family,

$$f(x|\theta) = h(x)\, c(\theta) \exp[w(\theta)\, t(x)].$$

Let $\eta = w(\theta)$ and let Ω be the parameter space (range) of η. If Ω is an open interval (a,b) $(a=-\infty, b=\infty$ possible), then X is from a one parameter regular exponential family 1PREF.

48) For a two parameter exponential family $\underline{\theta} = (\theta_1, \theta_2)$, and

$$f(x|\underline{\theta}) = h(x)\, c(\underline{\theta}) \exp[w_1(\underline{\theta}) t_1(x) + w_2(\underline{\theta}) t_2(x)].$$

Let $\eta_1 = w_1(\underline{\theta})$ and $\eta_2 = w_2(\underline{\theta})$.

If i) the parameter space of (η_1, η_2) $= \Omega =$ cross product of 2 open intervals,

ii) Neither η_1 and η_2 nor $t_1(x)$ and $t_2(x)$ satisfy a linearity constraint,

then X is from a 2 parameter regular exponential family 2PREF.

A linearity constraint is satisfied by $d_i(x) = t_i(x)$ or $d_i(x) = \eta_i$ $d_1(x)$ and $d_2(x)$ if $a_1 d_1(x) + a_2 d_2(x) = c$ for some constants a_1 and a_2 and for all x (or η_i) where at least one of the $a_i \neq 0$.

If $a_1 d_1(x) + a_2 d_2(x) = c$ $\forall x$ ($\forall n_i$)
iff $a_1 = a_2 = 0$, then $d_1(x)$ and $d_2(x)$
do not satisfy a linearity constraint.

49] Suppose X is from a 2PREF. If one of the parameters is known, then X is from a 1PREF.

50] All of the brand name exponential family distributions are regular distribution except the inverse Gaussian distribution.

1PREFS: $EXP(\theta)$, $Pois(\lambda)$, $bin(q, m)$ m known, single parameter pareto θ known, Weibull τ known, inverse exponential, two parameter pareto θ known, inverse Weibull τ known, geometric(β), negative binomial r known.

2PREFS; $N(\mu, \sigma^2)$, $LN(\mu, \sigma)$, $Gamma(\alpha, \beta)$, $beta(a, b)$.

ex] $X \sim U(0, \theta)$ is not a 1 parameter exponential family since the support $[0, \theta]$ depends on θ.

ex] $X \sim EXP(\theta)$ has $f(x|\theta) = \frac{1}{\theta} e^{-\frac{x}{\theta}}$, $x > 0$

$$= \underbrace{\frac{1}{\theta}}_{c(\theta)} \underbrace{\overset{I(x>0)\text{ better}}{(1)}}_{h(x)} \exp\left[\underbrace{-\frac{1}{\theta}}_{w(\theta)} \underbrace{x}_{t(x)}\right]$$

$\eta = -\frac{1}{\theta}$ has $\Omega = (-\infty, 0)$ since $\theta > 0$.

ex] $X \sim \text{pareto}(\alpha, \theta)$, θ known has

$$f(x|\alpha) = \alpha \theta^{\alpha} (\theta + x)^{-(\alpha+1)}, \quad x > 0$$

$$= \underbrace{\alpha \theta^{\alpha}}_{c(\alpha)} \underbrace{(1)}_{h(x)} \exp\left[\underbrace{-(\alpha+1)}_{w(\alpha)} \underbrace{\ln(\theta + x)}_{t(x)}\right]$$

$I(x>0)$ better

$\eta = -(\alpha+1)$, $\Omega = (-\infty, -1)$ since $\alpha > 0$.

ex] $X \sim N(\mu, \sigma^2)$ $\quad \underline{\theta} = (\mu, \sigma^2)$

$$f(x) = \underbrace{\frac{1}{\sqrt{2\pi}}}_{h(x)} \underbrace{\frac{1}{\sigma} e^{-\frac{\mu^2}{2\sigma^2}}}_{c(\underline{\theta})} \exp\left[\underbrace{\frac{\mu}{\sigma^2}}_{w_1(\underline{\theta})} \underbrace{x}_{t_1(x)} \underbrace{- \frac{1}{2\sigma^2}}_{w_2(\underline{\theta})} \underbrace{x^2}_{t_2(x)}\right] \quad (*)$$

$\eta_1 = \frac{\mu}{\sigma^2}$, $\eta_2 = \frac{-1}{2\sigma^2}$, $\Omega = (-\infty,\infty) \times (-\infty,0)$.

However if μ is known,

$\eta_1 = -2\mu\eta_2$, and η_1 and η_2 satisfy

a linearity constraint.

use $f(x) = \underbrace{\frac{1}{\sqrt{2\pi}}}_{h(x)} \underbrace{\frac{1}{\sigma}}_{c(\sigma^2)} \exp\left[\underbrace{\frac{-1}{2\sigma^2}}_{w(\sigma^2)} \underbrace{(x-\mu)^2}_{t(x)}\right]$.

51] p394 Let $X_1, \ldots, X_n$ be iid with pdf or pmf $f(x|\theta)$. Then the joint pdf or pmf is $f(x_1,\ldots,x_n|\theta) = f(\underset{\sim}{x}|\theta)$.

(Fisher's) information (number) =

$I_n(\theta) = I(\theta) = E_\theta\left(\left[\frac{d}{d\theta} \ln(f(\underset{\sim}{X}|\theta))\right]^2\right)$

$I_1(\theta) = E_\theta\left(\left[\frac{d}{d\theta} \ln f(X|\theta)\right]^2\right)$.

52] * If $X_1, \ldots, X_n$ are iid from a 1PREF,

then $I_1(\theta) = -E\left[\frac{d^2}{d\theta^2} \ln f(X|\theta)\right]$

and $I_n(\theta) = n\, I_1(\theta)$. see E2rev 73).

25.5

53] If $X_1, \ldots, X_n$ are iid from a 1PREF, let $\hat{\theta}_n$ be the MLE of θ.

Then $\sqrt{n}\,(\hat{\theta}_n - \theta) \xrightarrow{D} N\left(0, \frac{1}{I_1(\theta)}\right)$

and if $\tau'(\theta) \neq 0$, asymptotically efficient

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

So $\hat{\theta}_n \sim AN\left(\theta, \frac{1}{nI_1(\theta)}\right)$ and

$$\tau(\hat{\theta}_n) \sim AN\left(\tau(\theta), \frac{[\tau'(\theta)]^2}{nI_1(\theta)}\right).$$

54] Let matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ where

in this class usually $c = b$. $A_{11} = a$, $A_{12} = b$, $A_{21} = c$ $A_{22} = d$.

Then $A^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.

55] Let $L(\underline{\theta}) = L(\underline{\theta} \mid x_1, \ldots, x_n) = L(\underline{\theta} \mid \underline{x}) = f(x_1, \ldots, x_n \mid \underline{\theta}) = f(\underline{x} \mid \underline{\theta})$

with $\underline{\theta} = (\theta_1, \ldots, \theta_k)$ where usually $k = 2$.

The information matrix

$\underset{n}{I}(\underset{\sim}{\theta}) = [I_{ij}]$ where

$$I_{ij} = E\left[\frac{\partial}{\partial\theta_i} \ln f(\underset{\sim}{x}|\underset{\sim}{\theta}) \frac{\partial}{\partial\theta_j} \ln f(\underset{\sim}{x}|\underset{\sim}{\theta})\right].$$

If $X_1, \ldots, X_n$ are iid from a k-parameter exponential family, $I_{ij} = -E\left[\frac{\partial^2}{\partial\theta_i\partial\theta_j} \ln f(\underset{\sim}{x}|\underset{\sim}{\theta})\right]$.

If $\hat{\underset{\sim}{\theta}}$ is the MLE, then under regularity conditions,

$$\sqrt{n}(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta}) \xrightarrow{D} N_k\left[\underset{\sim}{0}, I_1^{-1}(\underset{\sim}{\theta})\right].$$

56) $\underset{\sim}{Y} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_k \end{pmatrix} \sim N_k(\underset{\sim}{\mu}, \Sigma) =$

$$N_k\left[\begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_k \end{pmatrix}, \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1k} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2k} \\ \vdots & \vdots & & \vdots \\ \sigma_{k1} & \sigma_{k2} & \cdots & \sigma_{kk} \end{pmatrix}\right],$$

a k-dimensional multivariate normal distribution, iff $\underset{\sim}{t}^T \underset{\sim}{Y}$ has a univariate normal distribution for any $k \times 1$ constant vector $\underset{\sim}{t}$. If Σ is positive definite, the pdf of $\underset{\sim}{Y}$ is

$$f(\underset{\sim}{z}) = \frac{1}{(2\pi)^{\frac{k}{2}}} \frac{1}{|\Sigma|^{\frac{1}{2}}} \exp\left[-\frac{1}{2} (\underset{\sim}{z}-\underset{\sim}{\mu})^T \Sigma^{-1} (\underset{\sim}{z}-\underset{\sim}{\mu})\right],$$

where $|\Sigma| = \det(\Sigma)$.

$$E(\underset{\sim}{Y}) = \begin{pmatrix} E(Y_1) \\ \vdots \\ E(Y_k) \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_k \end{pmatrix} = \underset{\sim}{\mu}.$$

The covariance matrix of $\underset{\sim}{Y}$ =

$\operatorname{cov}(\underset{\sim}{Y}) = \Sigma = [\sigma_{ij}]$ where

$\sigma_{ij} = \operatorname{cov}(Y_i, Y_j)$. Note that

$\sigma_{ii} = \operatorname{Var}(Y_i)$.

57] The most general definition of the information matrix is

$I_n(\underline{\theta}) = (I_{ij})$ where

$1 \times k$

$$I_{ij} = E\left[\frac{\partial}{\partial\theta_i}\ln(L(\underline{\theta}))\ \frac{\partial}{\partial\theta_j}\ln(L(\underline{\theta}))\right]$$

$$= E\left[\frac{\partial}{\partial\theta_i}\ell(\underline{\theta})\ \frac{\partial}{\partial\theta_j}\ell(\underline{\theta})\right]$$

where $\ell(\underline{\theta}) = \ln(L(\underline{\theta}))$ and

$$L(\underline{\theta}) = L(\underline{\theta}\,|\,\underline{X}) = L(\underline{\theta}\,|\,\underbrace{(X_1,\ldots,X_n)}_{\text{RVs}}).$$

When there is $k=1$ parameter,

$$I_n(\theta) = E\left(\left[\frac{d}{d\theta}\ln(L(\theta\,|\,\underline{X}))\right]^2\right).$$

When there are $k=2$ parameters, $I_n(\underline{\theta}) =$

$$\begin{pmatrix} E\left(\left[\frac{\partial}{\partial\theta_1}\ln L(\underline{\theta}\,|\,\underline{X})\right]^2\right) & E\left[\frac{\partial}{\partial\theta_1}\ln L(\underline{\theta}\,|\,\underline{X})\,\frac{\partial}{\partial\theta_2}\ln L(\underline{\theta}\,|\,\underline{X})\right] \\ & E\left(\left[\frac{\partial}{\partial\theta_2}\ln L(\underline{\theta}\,|\,\underline{X})\right]^2\right) \end{pmatrix}.$$

symmetric matrix

27.5

58] If $X_1, \ldots, X_n$ are iid from a k parameter exponential family, then

$$I_{ij} = -E\left[\frac{\partial^2}{\partial\theta_i \partial\theta_j} \ln L(\underline{\theta} \mid \underline{x})\right].$$

If $k = 1$, then

$$I_n(\theta) = -E\left[\frac{d^2}{d\theta^2} \ln L(\theta \mid \underline{x})\right]$$

$$= n I_1(\theta) \quad \text{where}$$

$$I_1(\theta) = -E\left[\frac{d^2}{d\theta^2} \ln L(\theta \mid \underset{\substack{\uparrow \\ \text{RV}, n=1}}{x})\right] \quad \text{as in 52]}$$

where $L(\theta \mid \underline{x}) = f(x_1, \ldots, x_n \mid \theta)$

and $L(\theta \mid x) = f(x \mid \theta)$.

redundant

Hence $I_1(\theta) = -E\left[\frac{d^2}{d\theta^2} \ln(f(x \mid \theta))\right]$ ✓

as in 52].

If $k = 2$, then $I_n(\underline{\theta}) =$

$$= -\begin{pmatrix} E\left[\frac{\partial^2}{\partial\theta_1^2}\ln L(\underline{\theta}|\underline{x})\right] & E\left[\frac{\partial^2}{\partial\theta_1\partial\theta_2}\ln L(\underline{\theta}|\underline{x})\right] \\ & E\left[\frac{\partial^2}{\partial\theta_2^2}\ln L(\underline{\theta}|\underline{x})\right] \end{pmatrix}.$$

symmetric

59) Formulas like

$$I_n(\theta) = -E\left[\frac{d^2}{d\theta^2}\ell(\theta)\right] = E\left(\left[\frac{d}{d\theta}\ell(\theta)\right]^2\right)$$

and

$$I_n(\underline{\theta}) = -\begin{bmatrix} E\left[\frac{\partial^2}{\partial\theta_1^2}\ell(\underline{\theta})\right] & E\left[\frac{\partial^2}{\partial\theta_1\partial\theta_2}\ell(\underline{\theta})\right] \\ & E\left[\frac{\partial^2}{\partial\theta_2^2}\ell(\underline{\theta})\right] \end{bmatrix}$$

are often given, but tend not to hold if $X_1,\ldots,X_n$ are not from a 1 or 2 parameter exponential family. For example, if $X_1,\ldots,X_n$ are iid $U(0,\theta)$, need to use 57) not 58).

60] Assume the 2nd partial derivative formulas work on "Exam C" questions (FAM-S ASTAM) unless told otherwise (or if $x_1, \dots, x_n$ are iid $U(0, \theta)$).

61] $V(\hat{\theta}) \approx \underbrace{Var(\hat{\theta})}_{\text{asymptotic variance}}$ = asymptotic variance of $\hat{\theta} = \dfrac{1}{n I_1(\theta)}$; AVAR$(\hat{\theta})$ is better

$= \dfrac{1}{I_n(\theta)}$.

asymptotic variance of $\tau(\hat{\theta}) =$

$V(\tau(\hat{\theta})) \approx$

$Var(\tau(\hat{\theta})) = \dfrac{[\tau'(\theta)]^2}{n I_1(\theta)} = \dfrac{[\tau'(\theta)]^2}{I_n(\theta)}$ if $\tau'(\theta) \neq 0$.

has form $\dfrac{\text{constant}}{n}$

↓

asymptotic variance of MLE

62] dist		$I_1(\theta)$	$\underbrace{Var(\hat{\theta})}$ ← AVAR$(\hat{\theta})$ better
$EXP(\theta)$		$\dfrac{1}{\theta^2}$	$\dfrac{\theta^2}{n}$
lognormal $LN(\mu, \sigma)$	μ	$\dfrac{1}{\sigma^2}$	$\dfrac{\sigma^2}{n}$
	σ	$\dfrac{2}{\sigma^2}$	$\dfrac{\sigma^2}{2n}$
pareto(α, θ)	α fixed	$I_1(\theta) = \dfrac{\alpha}{(\alpha+2)\theta^2}$	$\dfrac{(\alpha+2)\theta^2}{n\alpha}$
	θ fixed	$I_1(\alpha) = \dfrac{1}{\alpha^2}$	$\dfrac{\alpha^2}{n}$

← $\overbrace{Cov(\hat{\mu}, \hat{\sigma})}^{\text{asymptotic}} = 0$