2) $Y_1, \ldots, Y_n$ independent rvs ; $Y_i \sim Exp(\mu_i)$; $\mu_i = \beta x_i$

$x_1, \ldots, x_n$ are known positive constants not all $= 1$ and $\beta > 0$ unknown

(a) Find MLE of β, $\hat{\beta}$, and find large sample dist $\sqrt{n}(\hat{\beta} - \beta)$.

First, we write the likelihood for Y:

$$L = \prod_{i=1}^{n} \frac{1}{\beta x_i} \exp(-y_i/\beta x_i) \Rightarrow \text{log-likelihood: } \ell = \sum_{i=1}^{n} -\log(\beta x_i) - \frac{y_i}{\beta x_i}$$

$$\frac{d\ell}{d\beta} = \sum_{i=1}^{n} -\frac{x_i}{\beta x_i} + \frac{y_i x_i}{\beta^2 x_i^2} \overset{set}{=} 0 \Rightarrow \frac{n}{\hat{\beta}} = \sum_{i=1}^{n} \frac{y_i}{\hat{\beta}^2 x_i}$$

$$\Rightarrow n\hat{\beta} = \sum_{i=1}^{n} y_i/x_i \Rightarrow \boxed{\hat{\beta} = \frac{1}{n}\sum_{i=1}^{n} y_i/x_i}$$

This distribution is a member of the exponential family since x_i are known constants, so we know regularity conditions hold.

Thus, by large-sample properties of MLE, we know:

$$\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{d} N(0, I(\beta)^{-1})$$

We find $I(\beta) = \lim_{n\to\infty} \frac{1}{n} E\left[-\frac{d^2\ell}{d\beta^2}\right]$

$$\hookrightarrow \frac{d^2\ell}{d\beta^2} = \sum_{i=1}^{n} \frac{1}{\beta^2} - \frac{y_i x_i (2\beta x_i^2)}{\beta^4 x_i^4} = \sum_{i=1}^{n} \frac{1}{\beta^2} - \frac{2y_i}{\beta^3 x_i}$$

Then, $-E\left[\frac{d^2\ell}{d\beta^2}\right] = -\frac{n}{\beta^2} + \frac{2}{\beta^3}\sum_{i=1}^{n} E(y_i/x_i)$

Remember, x_i is a known constant, so $E(y_i/x_i) = \frac{1}{x_i}E(y_i) = \frac{1}{x_i}(\beta x_i) = \beta$

Thus, $-E\left[\frac{d^2\ell}{d\beta^2}\right] = \frac{-n}{\beta^2} + \frac{2}{\beta^3}\sum_{i=1}^{n}\beta = \frac{-n}{\beta^2} + \frac{2n\beta}{\beta^3} = \frac{-n+2n}{\beta^2} = \frac{n}{\beta^2}$

Hence, $I(\beta) = \lim_{n\to\infty}\frac{1}{n}\left(\frac{n}{\beta^2}\right) = \frac{1}{\beta^2}$, and so $I(\beta)^{-1} = \beta^2$

Thus, $\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{d} N(0, \beta^2)$

(b) Find a pivotal quantity and use it to construct an <u>exact</u> 95% CI for β

In exponential family form, our likelihood is:

$$L = \exp\left\{-\sum_{i=1}^{n}\left(\frac{y_i}{\beta x_i}\right) + \log\left(\prod_{i=1}^{n}\frac{1}{\beta x_i}\right)\right\} = \exp\left\{-\sum_{i=1}^{n}\left[\left(y_i/\beta x_i\right) - \log(\beta x_i)\right]\right\}$$

From the exponential family form, we see that $\sum_{i=1}^{n} y_i/x_i$ is a sufficient statistic, so we will investigate with this. for $(\frac{1}{\beta})$ and hence for β.

We will find the distribution of $\sum_{i=1}^{n} y_i/x_i$ to see if it is parameter free.

Let $Z_i = y_i/x_i \Rightarrow Y_i = Z_i X_i \Rightarrow \frac{dY_i}{dZ_i} = X_i$

So, $f_{Z_i}(z_i) = f_{Y_i}(z_i x_i)\cdot \underset{\text{known}, >0}{x_i} = \frac{1}{\beta x_i}\exp\left(-z_i x_i/\beta x_i\right)\cdot x_i = \frac{1}{\beta}\exp\left(-z_i/\beta\right)$,

So, $Z_i = \frac{y_i}{x_i} \sim \text{Exp}(\beta) \Rightarrow \sum_{i=1}^{n}\frac{y_i}{x_i} \sim \text{Gamma}(n,\beta)$

Thus, $\frac{1}{\beta}\sum_{i=1}^{n}\frac{y_i}{x_i} \sim \text{Gamma}(n,1)$, which is a distribution $\perp\!\!\!\perp$ of β

$\Rightarrow \frac{1}{\beta}\sum_{i=1}^{n}\frac{y_i}{x_i}$ is a pivotal quantity.

Thus, a exact 95% CI can be constructed of the form $a \le \frac{1}{\beta}\sum_{i=1}^{n} y_i/x_i \le b$,

where a is the $\frac{\alpha}{2}$ quantile of the Gamma$(n,1)$ distribution,

so $F_{\Gamma(n,1)}(a) = \frac{\alpha}{2}$, so let $a = F^{-1}_{\Gamma(n,1)}(\alpha/2)$

and b is the $1-\frac{\alpha}{2}$ percentile of Gamma$(n,1)$ distribution,

So similarly, let $b = F^{-1}_{\Gamma(n,1)}(1-\frac{\alpha}{2})$

$\alpha = 0.05$ here, so $\frac{\alpha}{2} = 0.025$ and $1-\frac{\alpha}{2} = 0.975$

Thus, an exact 95% CI for β is: $\left\{\beta : a \le \frac{1}{\beta}\sum_{i=1}^{n} y_i/x_i \le b\right\} = \left\{\beta : \frac{\sum_{i=1}^{n} y_i/x_i}{b} \le \beta \le \frac{\sum_{i=1}^{n} y_i/x_i}{a}\right\}$

$$\Rightarrow \left\{\beta : \frac{\sum_{i=1}^{n}(y_i/x_i)}{F^{-1}_{\Gamma(n,1)}(0.975)} \le \beta \le \frac{\sum_{i=1}^{n}(y_i/x_i)}{F^{-1}_{\Gamma(n,1)}(0.025)}\right\}$$

Could put $n\hat{\beta}$ in place of $\sum(y_i/x_i)$ too

(c) Estimator of β : $\tilde{\beta} = \dfrac{\sum_{i=1}^{n} Y_i}{\sum_{i=1}^{n} X_i}$. Show the finite sample efficiency of $\tilde{\beta}$ relative to $\hat{\beta} \leq 1$

- The efficiency of $\tilde{\beta}$ relative to $\hat{\beta}$ is $\dfrac{E[(\hat{\beta}-\beta)^2]}{E[(\tilde{\beta}-\beta)^2]}$; just comparing MSE !

  Also, MSE = Variance + BIAS^2

  <u>First, focusing on $\hat{\beta}$:</u>

  $\hat{\beta} = \frac{1}{n}\sum_{i=1}^{n} Y_i/X_i$; $E(\hat{\beta}) = \frac{1}{n}\sum_{i=1}^{n} E(Y_i/X_i)$; in (b) we showed $\frac{Y_i}{X_i} \sim Exp(\beta)$

  $= \beta \Rightarrow \text{BIAS} = 0$ So $E(Y_i/X_i) = \beta$

  and $Var(\hat{\beta}) = \frac{1}{n^2}\sum_{i=1}^{n} Var(Y_i/X_i) = \frac{1}{n^2}\sum_{i=1}^{n}\beta^2 = \frac{\beta^2}{n}$

  Thus, $E[(\hat{\beta}-\beta)^2] = \beta^2 + 0 = \frac{\beta^2}{n}$

- <u>Now, focusing on $\tilde{\beta}$:</u>

  $\tilde{\beta} = \dfrac{\sum_{i=1}^{n} Y_i}{\sum_{i=1}^{n} X_i}$ constants. $\rightarrow E(\tilde{\beta}) = \dfrac{\sum_{i=1}^{n} E(Y_i)}{\sum_{i=1}^{n} X_i} = \dfrac{\sum_{i=1}^{n} \beta X_i}{\sum_{i=1}^{n} X_i} = \beta$, so no <u>BIAS</u>

  $Var(\tilde{\beta}) = \dfrac{\sum_{i=1}^{n} Var(Y_i)}{\left(\sum_{i=1}^{n} X_i\right)^2} = \dfrac{\sum_{i=1}^{n} \beta^2 X_i^2}{\left(\sum_{i=1}^{n} X_i\right)^2} = \beta^2\left(\dfrac{\sum_{i=1}^{n} X_i^2}{\left(\sum_{i=1}^{n} X_i\right)^2}\right)$

  Thus, $E[(\tilde{\beta}-\beta)^2] = \beta^2\left(\dfrac{\sum_{i=1}^{n} X_i^2}{\left(\sum_{i=1}^{n} X_i\right)^2}\right)$

<u>Note</u> : Using Holder's Inequality, $\left(\sum_{i=1}^{n} X_i\right)^2 = \left(\sum_{i=1}^{n}(1\cdot X_i)\right)^2 \leq \underbrace{\left(\sum_{i=1}^{n} 1^2\right)}_{n} \cdot \sum_{i=1}^{n} X_i^2$

$\Rightarrow \left(\sum_{i=1}^{n} X_i\right)^2 \leq n\sum_{i=1}^{n} X_i^2 \Rightarrow \left(\dfrac{\sum_{i=1}^{n} X_i^2}{\left(\sum_{i=1}^{n} X_i\right)^2}\right) \geq \frac{1}{n}$

- Thus, the finite sample efficiency of $\tilde{\beta}$ relative to $\hat{\beta}$ is :

  $\dfrac{\frac{1}{n}\beta^2}{\beta^2\left(\sum_{i=1}^{n} X_i^2 / \left(\sum_{i=1}^{n} X_i\right)^2\right)} \leq \dfrac{\frac{1}{n}\beta^2}{\beta^2 \cdot \frac{1}{n}} = 1$ ✓

(d) Now different mean: $\frac{1}{\mu_i} = \alpha + \gamma x_i$; α & γ unknown parameters

Find a minimal SS for (α, γ).

Our original function is $f_{Y_i}(y_i) = \frac{1}{\mu_i} \exp\left(-y_i/\mu_i\right)$

Now, $\frac{1}{\mu_i} = \alpha + \gamma x_i \Rightarrow f_{Y_i}(y_i) = (\alpha + \gamma x_i) \exp\left(-y_i(\alpha + \gamma x_i)\right)$

The likelihood is: $L = \prod_{i=1}^{n} (\alpha + \gamma x_i) \exp\left(-y_i(\alpha + \gamma x_i)\right)$

$\Rightarrow$ log-likelihood: $\ell = \sum_{i=1}^{n} \left[\log(\alpha + \gamma x_i) - y_i(\alpha + \gamma x_i)\right]$

Thus, our likelihood in exponential family form is:

$$L = \exp\left\{\sum_{i=1}^{n} -y_i\alpha - y_i x_i \gamma + \log(\alpha + \gamma x_i)\right\}$$

So, we have $\theta = (-\alpha, -\gamma)$, $T(x,y) = \left(+\sum_{i=1}^{n} y_i, +\sum_{i=1}^{n} y_i x_i\right)$, $b(\theta) = -\sum_{i=1}^{n} \log(\alpha + \gamma x_i)$

Thus, a minimal sufficient statistic for (α, γ) is

$$\left(\sum_{i=1}^{n} y_i, \sum_{i=1}^{n} y_i x_i\right) \leftarrow \text{Nor redundant}$$

based on the exponential family form of rank 2.

Proof of Minimal:

$$\frac{\log f_{Y|\alpha,\gamma}}{\log f_{Z|\alpha,\gamma}} \propto \frac{\sum_{i=1}^{n} -y_i\alpha - y_i\gamma x_i}{\sum_{i=1}^{n} -z_i\alpha - z_i\gamma x_i} = \frac{-\alpha\sum_{i=1}^{n} y_i - \gamma\sum_{i=1}^{n} y_i x_i}{-\alpha\sum_{i=1}^{n} z_i - \gamma\sum_{i=1}^{n} z_i x_i}$$

does not depend on (α, γ) iff $\sum y_i = \sum z_i$ and $\sum y_i x_i = \sum z_i x_i$,

so we have minimal sufficient stat!

(e) By appropriate conditioning, obtain the conditional score function for $\underline{\underline{\gamma}}$? (eliminating α).

You don't need to simplify.

From (d), we have the full likelihood : $L = \prod_{i=1}^{n} (\alpha + \gamma x_i) \exp(-y_i(\alpha+\gamma x_i))$

and we found $\sum_{i=1}^{n} Y_i$ is sufficient for α and $\sum_{i=1}^{n} Y_i X_i$ is sufficient for γ.

Now, we need to find these distributions.

· We know $Y_i \sim Exp(\alpha + \gamma x_i)$, not iid!

↳ Now, let's look at $Y_i X_i$:

$$\text{Let } Z_i = Y_i X_i \Rightarrow Y_i = \frac{X_i}{Z_i}, \text{ so } \frac{dY_i}{dZ_i} = \frac{-X_i}{Z_i^2} \Rightarrow f_{Z_i} = \underbrace{(\alpha+\gamma x_i) \exp\left(\frac{-x_i}{Z_i}(\alpha+\gamma x_i)\right) \cdot \frac{X_i}{Z_i^2}}_{\text{Does not look familiar}}$$

Maybe inverse gamma?

So $Y_i X_i \sim \text{Inv Gamma}\left(1, x_i(\alpha+\gamma x_i)\right)$ (also not iid)!

Yes! $\frac{\left(x_i(\alpha+\gamma x_i)\right)^1}{\Gamma(1)} Z^{-1-1} \exp\left\{\frac{-x_i(\alpha+\gamma x_i)}{Z_i}\right\}$

First, $Y_i \sim Exp\left(\frac{1}{\alpha+\gamma x_i}\right)$

$$\Rightarrow Y_i(\alpha+\gamma x_i) \sim Exp(1) \Rightarrow \sum_{i=1}^{n} Y_i(\alpha+\gamma x_i) \sim Gamma(n,1)$$

$$\Rightarrow \alpha \sum_{i=1}^{n} Y_i + \gamma \sum_{i=1}^{n} Y_i x_i \sim Gamma(n,1)$$

Unfortunately, finding explicit expression for $\sum_{i=1}^{n} Y_i$ is not feasible.

So, the Conditional likelihood given $\sum_{i=1}^{n} Y_i$ is :

$$L_c(\gamma) = \frac{\exp\left\{\cancel{-\alpha \sum_{i=1}^{n} Y_i} - \gamma \sum_{i=1}^{n} Y_i x_i + \cancel{\sum_{i=1}^{n} \log(\alpha+\gamma x_i)}\right\}}{\sum_{\tilde{y} \in S} \left[\exp\left\{\cancel{-\alpha \sum_{i=1}^{n} \tilde{Y}_i} - \gamma \sum_{i=1}^{n} \tilde{Y}_i x_i + \cancel{\sum_{i=1}^{n} \log(\alpha+\gamma x_i)}\right\}\right]}, \text{ where } S = \left\{Y : \sum_{i=1}^{n} Y_i = \sum_{i=1}^{n} Y_i\right\}$$

$$= \frac{\exp\left\{-\gamma \sum_{i=1}^{n} Y_i x_i\right\}}{\sum_{\tilde{y} \in S} \exp\left\{-\gamma \sum_{i=1}^{n} \tilde{Y}_i x_i\right\}}$$

So, $\ell_c(\gamma) = -\gamma \sum_{i=1}^{n} y_i x_i - \log\left[\sum_{\tilde{y} \in S} \exp\left\{-\gamma \sum_{i=1}^{n} \tilde{y}_i x_i\right\}\right]$

→ Conditional score equation: $\frac{\partial \ell_c(\gamma)}{\partial \gamma} = -\sum_{i=1}^{n} y_i x_i - \frac{\sum_{\tilde{y} \in S} \left(\sum_{i=1}^{n} \tilde{y}_i x_i\right) \exp\left\{-\gamma \sum_{i=1}^{n} \tilde{y}_i x_i\right\}}{\sum_{\tilde{y} \in S} \exp\left\{-\gamma \sum_{i=1}^{n} \tilde{y}_i x_i\right\}}$