Name: Dat Quoc Ngo

NET-ID: dqn170000

Course: CS 6364.002

Homework 3

1: MLE_questions #1

Likelihood $L(\theta) = \prod_{i=1}^{N} \theta x_i^{\theta-1}$

Taking natural logarithm

$$\ln L(\theta) = \ln \left(\sum_{i=1}^{N} \theta x_i^{\theta-1} \right)$$

$$= \sum_{i=1}^{N} \ln(\theta) + (\theta - 1) \sum_{i=1}^{N} \ln(x_i)$$

$$\frac{dL}{d\theta} = \frac{N}{\theta} + \sum_{i=1}^{N} \ln(x_i) = 0$$

$$\Rightarrow \frac{\theta}{N} = -\left(\sum_{i=1}^{N} \ln(x_i)\right)^{-1}$$

$$= -\left[\ln(0.4) + \ln(0.7) + \ln(0.9) + \ln(0.6) + \ln(0.5) + \ln(0.7)\right]^{-1} \times 6$$

$$= 0.34025 \times 6 = 2.0415$$

(2) MLE - questions #2

Given uniform distribution

$$0 \leq x_i \leq \theta$$

$$\Rightarrow f(x_i) = \frac{1}{\theta - 0} = \frac{1}{\theta}$$

Log likelihood

$$L(\theta) = \prod_{i=1}^{N} \frac{1}{\theta} = \theta^{-N}$$

Taking natural logarithm

$$\ln L(\theta) = -n \ln(\theta)$$

$$\frac{dL}{d\theta} = -n\frac{1}{\theta} < 0 \text{ given } \theta < 0$$

Hence $L(\theta)$ is decreasing and it's maximized at $\theta = X_N$

that X_N is the largest value on the interval $(0, \theta)$

$$\boxed{\theta = X_N}$$

③ MLE - questions #3

$$f(x) = \frac{1}{\beta} e^{-\frac{x}{\beta}} \quad \text{with} \quad 0 < x < \infty$$

Log likelihood

$$L(\beta) = \prod_{i=1}^{N} \frac{1}{\beta} e^{-\frac{x_i}{\beta}}$$

Taking natural logarithm

$$\ln L(\beta) = \sum_{i=1}^{N} \ln\left(\frac{1}{\beta} \cdot e^{-\frac{x_i}{\beta}}\right)$$

$$= \sum_{i=1}^{N} \ln\left(\frac{1}{\beta}\right) + \sum_{i=1}^{N} \ln e^{-x_i/\beta}$$

$$= N \ln\frac{1}{\beta} - \frac{1}{\beta} \cdot \sum_{i=1}^{N} x_i$$

$$\frac{dL}{d\beta} = N.\beta + \frac{1}{\beta^2} \sum_{i=1}^{N} x_i = 0$$

$$\Rightarrow \frac{1}{\beta^2} \left(N.\beta^3 + \sum_{i=1}^{N} x_i \right) = 0$$

Since $\frac{1}{\beta^2} > 0$

$$\Rightarrow N.\beta^3 + \sum_{i=1}^{N} x_i = 0$$

$$\Rightarrow \beta^3 = -\frac{1}{N} \sum_{i=1}^{N} x_i$$

$$\boxed{\Rightarrow \beta = \sqrt[3]{\frac{-1}{N} \cdot \sum_{i=1}^{N} x_i}}$$

(4) MLE - questions # 4

Since $0 \leq x < \infty$

$$\Rightarrow \Gamma(\alpha) = (\alpha - 1)!$$

Then

$$f(x) = \frac{1}{(\alpha-1)!} \lambda^{\alpha} x^{\alpha-1} e^{-\lambda x}$$

Log likelihood

$$L(\lambda, x_i) = \prod_{i=1}^{N} \frac{1}{(\alpha-1)!} \lambda^{\alpha} x_i^{\alpha-1} e^{-\lambda x_i}$$

Taking natural logarithm

$$\ln L(\lambda, x_i) =$$

$$\sum_{i=1}^{N} \ln \frac{1}{(\alpha-1)!} + \sum_{i=1}^{N} \ln \lambda^{\alpha} + \sum_{i=1}^{N} \ln (x_i^{\alpha-1}) + \sum_{i=1}^{N} \ln (e^{-\lambda x_i})$$

$$= -N \ln (\alpha-1)! + N \ln \lambda^{\alpha} + (\alpha-1) \sum_{i=1}^{N} \ln x_i - \lambda \sum_{i=1}^{N} x_i$$

Taking 1st derivative

$$\frac{dL}{d\lambda} = 0 + N \cdot \frac{1}{\lambda^{\alpha}} \alpha (\lambda^{\alpha-1}) + 0 - \lambda \sum_{i=1}^{N} x_i = 0$$

$$= \frac{N\alpha}{\lambda} - \lambda = 0$$

$$\Rightarrow \frac{N \cdot \alpha - \lambda^2}{\lambda} = 0$$

$\Rightarrow \lambda = N\alpha$ (1)

$\Rightarrow \lambda = \pm\sqrt{N\alpha}$ that $\lambda \neq 0$

Taking 2nd derivative

$$\frac{dL}{d\lambda} = \frac{N\alpha}{\lambda} - \lambda \Rightarrow \frac{d^2L}{d\lambda^2} = \frac{-N\alpha}{\lambda^2} - 1$$

From (1) $\Rightarrow \lambda^2 = N\alpha \Rightarrow \alpha = \frac{\lambda^2}{N}$ that $N > 0$

$\Rightarrow \alpha > 0$

hence $\frac{d^2L}{d\lambda^2} = \frac{-N\alpha}{\lambda^2} - 1 < 0$ given

$\lambda = \pm\sqrt{N\alpha}$.

$\Rightarrow$ MLE

$$\boxed{\lambda = \pm\sqrt{N\alpha} \text{ that } \alpha > 0}$$

(5) MLE-questions #5

$$f(x) = \frac{1}{2} e^{-|x-\theta|}$$

Log likelihood

$$L(\theta) = \prod_{i=1}^{N} \frac{1}{2} e^{-|x_i - \theta|}$$

$$\ln L(\theta) = \sum_{i=1}^{N} \ln \frac{1}{2} + \sum_{i=1}^{N} \ln e^{-|x_i - \theta|}$$

$$= -N \ln 2 - \sum_{i=1}^{N} |x_i - \theta|$$

Taking derivative

$$\frac{dL}{d\theta} = 0 - \sum_{i=1}^{N} \left(|x_i - \theta| \right)' = - \sum_{i=1}^{N} \left(\sqrt{(x_i - \theta)^2} \right)'$$

$$= - \sum_{i=1}^{N} \frac{1}{2} \frac{2 \cdot (x_i - \theta)}{\sqrt{(x_i - \theta)^2}} (-1)$$

$$= \sum_{i=1}^{N} \frac{(x_i - \theta)}{|x_i - \theta|} = 0$$

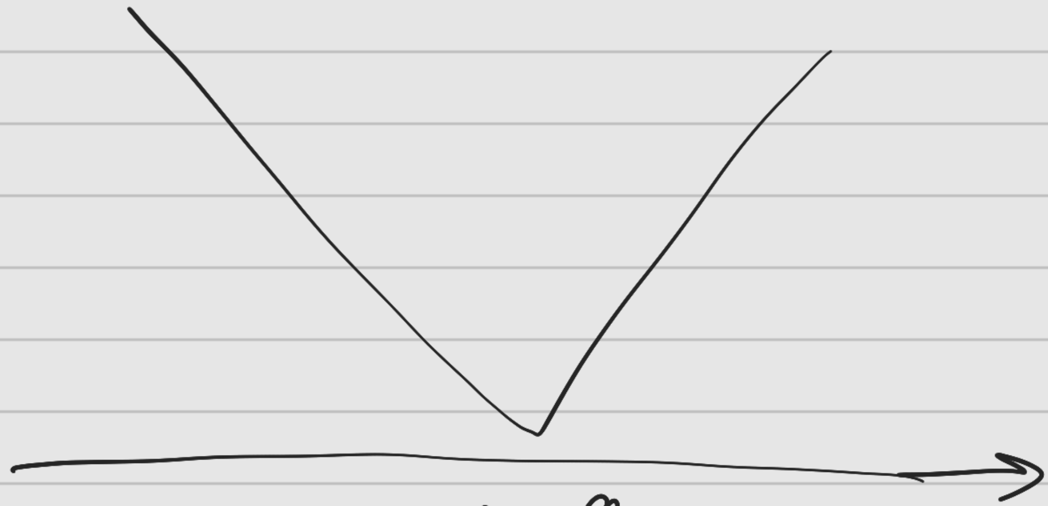
If $x_i < \theta$

$$\Rightarrow \frac{dL}{d\theta} = \sum_{i=1}^{N} \frac{(x_i - \theta)}{-(x_i - \theta)} = - \sum_{i=1}^{N} 1 = -N < 0$$

If $x_i > \theta$

$$\Rightarrow \frac{dL}{d\theta} = \sum_{i=1}^{N} \frac{(x_i - \theta)}{(x_i - \theta)} = \sum_{i=1}^{N} 1 = N > 0$$

We can plot & take $x_i = \theta$ as non-existing critical point



$x_i = \theta$

$x_i < \theta$

$$\Rightarrow \frac{dL}{d\theta} = -N < 0$$

$x_i > \theta$

$$\Rightarrow \frac{dL}{d\theta} = N > 0$$

Taking 2nd derivative

$$\frac{dL}{d\theta} = \sum_{i=1}^{N} \frac{x_i - \theta}{|x_i - \theta|} \Rightarrow \frac{d^2L}{d\theta^2} = \sum_{i=1}^{N} \frac{-1\,|x_i - \theta| - \frac{(x_i-\theta)\cdot(x_i-\theta)}{|x_i-\theta|}}{(x_i - \theta)^2}$$

$$= \sum_{i=1}^{N} \frac{-|x_i - \theta| - \frac{(x_i-\theta)^2}{|x_i - \theta|}}{(x_i - \theta)^2\,|x_i - \theta|} < 0 \text{ given}$$

$x_i \neq \theta$

. hence $x_i \neq \theta$

$\Rightarrow$ $\boxed{\text{MLE of } \theta \text{ when } \theta \neq x}$

⑥ PRML # 2.2 page 127

Distribution is normalized if sum of all results = 1

When

$x = -1$

$$\Rightarrow \sum_{\mu} p(-1|\mu) = \left(\frac{1-\mu}{2}\right)^{(1+1)/2} \cdot \left(\frac{1+\mu}{2}\right)^{(1-1)/2}$$

$$= \frac{1-\mu}{2}$$

When $x = 1$

$(1-1)/2 \qquad (1+1)/2$

$$\Rightarrow \sum_{\mu} p(1,\mu) = \left(\frac{1-\mu}{2}\right)^{0} \cdot \left(\frac{1+\mu}{2}\right)$$

$$= \frac{1+\mu}{2}$$

$$\Rightarrow \sum_{x}\sum_{\mu} p(x|\mu) = \left(\frac{1+\mu}{2}\right) + \left(\frac{1-\mu}{2}\right)$$

$$= \frac{1+1+\mu-\mu}{2} = 1$$

$$\Rightarrow \boxed{p(x|\mu) \text{ is normalized}}$$

Hence,

~~#~~ Mean $E(x) = \sum_{x \in \{-1,1\}} x . p(x|\mu)$

$$= (-1)p(x=-1|\mu) + 1\, p(x=1|\mu)$$

$$= -\left(\frac{1-\mu}{2}\right) + \frac{1+\mu}{2}$$

$$= -\frac{1}{2} + \frac{\mu}{2} + \frac{1}{2} + \frac{\mu}{2} = \mu$$

$\Rightarrow$ Mean $E(x) = \mu$

\# Variance

$$\mathrm{var}(x) = \sum_{x \in \{-1,1\}} (x - E(x))^2 p(x|\mu)$$

$$= \sum_{x \in \{-1,1\}} (x-\mu)^2 p(x|\mu)$$

$$= (-1-\mu)^2 p(-1|\mu)$$
$$+ (1-\mu)^2 p(1|\mu)$$

$$= (1+\mu)^2 \left(\frac{1-\mu}{2}\right) + (1-\mu)^2 \left(\frac{1+\mu}{2}\right)$$

$$= (1 + 2\mu + \mu^2)\left(\frac{1-\mu}{2}\right)$$
$$+ (1 - 2\mu + \mu^2)\left(\frac{1+\mu}{2}\right)$$

$$= (1+\mu^2)\left(\frac{1-\mu}{2} + \frac{1+\mu}{2}\right) + 2\mu\left(\frac{1-\mu}{2} - \frac{1+\mu}{2}\right)$$

$$= (1+\mu^2) + 2\mu\left(\frac{-2\mu}{2}\right) = 1 + \mu^2 - 2\mu^2 = 1 - \mu^2$$

$$\boxed{\Rightarrow \mathrm{Var}(x) = 1 - \mu^2}$$

$\rightarrow$ Entropy $H[x]$

$$H[x] = \sum_{x \in \{-1, 1\}} p(x|\mu) \ln p(x|\mu)$$

$$= p(-1|\mu) \ln p(-1|\mu) + p(1|u) \ln p(1|u)$$

$$= \left(\frac{1-u}{2}\right) \ln \left(\frac{1-u}{2}\right) + \left(\frac{1+u}{2}\right) \ln \left(\frac{1+u}{2}\right)$$

$$= \ln \left[\left(\frac{1-u}{2}\right)^{(1-u)/2} \cdot \left(\frac{1+u}{2}\right)^{(1+u)/2} \right]$$

⑦ PRML # 2.4 page 168

Set $p(m/\mu) = \sum_{m=0}^{N} \binom{N}{m} \mu^m (1-\mu)^{N-m} = 1$ ①

$$\frac{dp}{du} = \sum_{m=0}^{N} \binom{N}{m} \cdot \Big[m . \mu^{m-1} (1-\mu)^{N-m}$$

$$+ \mu^m . (N-m) (1-\mu)^{N-m-1} (-1) \Big] = 0$$

$$= \sum_{m=0}^{N} \binom{N}{m} \mu^{m-1} . (1-u)^{N-m-1} \left[m(1-\mu) - \mu (N-m) \right]$$

$$= \sum_{m=0}^{N} \binom{N}{m} \mu^{m-1} (1-\mu)^{N-m-1} . (m - \mu N) = 0$$ ②

$$\Rightarrow \quad m = N\mu \qquad (3)$$

Hence,

$$E(m) = \sum_{m=0}^{N} m \, \mathrm{Bin}(m|N,\mu)$$

$$= m \sum_{m=0}^{N} \binom{N}{m} \mu^m (1-\mu)^{N-m}$$

From (1) and (3), $E(m) = m \cdot 1 = N\mu$

From (2)

$$\frac{dp}{d\mu} = \sum_{m=0}^{N} \mu^{m-1} (1-\mu)^{N-m-1} (m - N\mu)$$

$$\frac{d^2p}{d\mu^2} = \sum_{m=0}^{N} \Big[(m-1)\mu^{m-2} (1-\mu)^{N-m-1} (m-N\mu)$$

$$+ \mu^{m-1} \Big((N-m-1)(1-\mu)^{N-m-2} (-1)(m-N\mu)$$

$$- N(1-\mu)^{N-m-1} \Big) \Big]$$

$$= \sum_{m=0}^{N} \Big[\mu^{m-2} (1-\mu)^{N-m-1} (m-1)(m-N\mu)$$

$$- \mu^{m-1} (1-\mu)^{N-m-2} (N-m-1)(m-N\mu)$$

$$- N(1-\mu)^{N-m-1} \mu^{m-1} \Big]$$

$$= \sum_{m=0}^{N} \mu^{m-2}(1-\mu)^{N-m-2} \cdot$$

$$\left[(1-\mu)(m-1)(m-N\mu) - \mu(N-m-1)(m-N\mu) - N(1-\mu)\mu \right] = 0$$

$$\Rightarrow (1-\mu)(m^2 - mN\mu - m + \cancel{N\mu} - \cancel{N\mu}) - \mu(N-m-1)(m-N\mu) = 0$$

$$\Rightarrow (1-\mu)(m^2 - mN\mu - m) - \mu(N-m-1)(m-N\mu) = 0$$

$$\Rightarrow (1-\mu)m(m-N\mu) - (1-\mu)m - \mu(N-m-1)(m-N\mu) = 0$$

$$\Rightarrow (m-N\mu)(m - \cancel{\mu m} - N\mu + \cancel{\mu m} + \mu) - (1-\mu)m = 0$$

$$\Rightarrow (m-N\mu)(m-N\mu+\mu) - (1-\mu)m = 0$$

$$\Rightarrow (m-N\mu)^2 + \mu m - N\mu^2 - m + \mu m = 0$$

$$\Rightarrow (m-N\mu)^2 = N\mu^2 + m - 2\mu m$$

From (3) $m = N\mu$

and $E(m) = N\mu$

$$\Rightarrow (m - E(m))^2 = N\mu^2 + N\mu - 2N\mu^2$$
$$= N\mu - N\mu^2$$
$$= N\mu(1-\mu)$$

Hence,

$var[m)$

$$\boxed{= \sum_{m=0}^{N} (m - E[m])^2 \overbrace{Bin(m|N,\mu)}^{=1}}$$

$$\boxed{= (m - E[m])^2 = N\mu(1-\mu)}$$

(8) PRML #2.6 page 128

Beta dist.

$$Beta(\mu|a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \mu^{a-1}(1-\mu)^{b-1} \quad (2.13)$$

$$\Gamma(x) \text{ is gamma function } \int_0^{\infty} u^{x-1} e^{-u} du \quad (1.141)$$

$$\int_0^1 Beta(\mu|a,b)\, d\mu = 1 \quad (2.14) \text{ normalized}$$

E...

From (2.265),

$$\int_0^1 \mu^{a-1}(1-\mu)^{b-1}\,d\mu = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

$$E[\mu] = \int_0^1 \mu \,\mathrm{Beta}(\mu \mid a,b)$$

$$= \int_0^1 \mu \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \cdot \mu^{a-1}(1-\mu)^{b-1}\,d\mu$$

$$= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \underbrace{\int_0^1 \mu^{a} \cdot (1-\mu)^{b-1}\,d\mu}_{1}$$

From 1

$$\int_0^1 \mu^{(a-1)+1} \cdot (1-\mu)^{b-1}\,d\mu = \frac{\Gamma(a+1)\,\Gamma(b)}{\Gamma(a+b+1)}$$

$$\Rightarrow \boxed{E[\mu] = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \cdot \frac{\Gamma(a+1)\Gamma(b)}{\Gamma(a+b+1)} = \frac{a}{a+b}}$$

using identity $\Gamma(t+1) = t\Gamma(t)$

$$Var[u] = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^{\infty} u^{a+2} (1-u)^{b-1} du - \left(\frac{a}{a+b}\right)^2$$

$$= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \cdot \frac{\Gamma(a+2)\,\Gamma(b)}{\Gamma(a+b+2)} - \left(\frac{a}{a+b}\right)^2$$

$$= \frac{a}{a+b} \cdot \frac{(a+1)}{(a+b+1)} - \left(\frac{a}{a+b}\right)^2$$

$$\Rightarrow Var[u] = \frac{a(a+1)}{(a+b)(a+b+1)} - \frac{a^2}{(a+b)^2}$$

$$= \frac{(a^2+a)(a+b) - a^2(a+b+1)}{(a+b)^2(a+b+1)}$$

$$= \frac{\cancel{a^3} + \cancel{ba^2} + \cancel{a^2} + ab - \cancel{a^3} - \cancel{a^2b} - \cancel{a^2}}{(a+b)^2(a+b+1)}$$

$$(a+b) \quad (a+b+1)$$

$$= \boxed{\frac{ab}{(a+b)^2(a+b+1)} = \mathrm{Var}[\mu]}$$

$$\frac{d}{d\mu} \mathrm{Beta}(\mu|a,b) = \frac{d}{d\mu} \cdot \mu^{a-1} \cdot (1-\mu)^{b-1}$$

$$= (a-1)\mu^{a-2}(1-\mu)^{b-1}$$
$$- (b-1)(1-\mu)^{b-2}\mu^{a-1} = 0$$

$$\Rightarrow \mu^{a-2}(1-\mu)^{b-2}\left[(a-1)(1-\mu)\right.$$

$$\left. -(b-1)\mu\right] = 0$$

$$\Rightarrow (a-1) - \mu(a+b-2) = 0$$

$$\boxed{\mu = \frac{a-1}{a+b-2} = \text{mode}}$$

(9) PRML #1 2.41 page 134

Ref. gama func 1.141 → gam

2.146 normalized

2.146

$$\mathrm{Gamma}(\lambda | a, b) = \frac{1}{\Gamma(a)} b^a \lambda^{a-1} \exp(-b\lambda)$$

Dister

Gamm fun $\Gamma(x) = \int_0^{\infty} u^{x-1} e^{-u} du$

Gama Distribution is normalized

if $\int_0^1 \mathrm{Gamma}(\lambda | a, b) = 1$

$$= \int_0^1 \frac{1}{\Gamma(a)} b^a \lambda^{a-1} e^{-b\lambda} = 1$$

$$\Rightarrow \int_0^1 b^a \cancel{\lambda}^{a-1} e^{-b\lambda} = \Gamma(a)$$

$$\Gamma(a) = \int_0^\infty x^{a-1} e^{-x} dx$$

Set $x = ub$ since $x > 0 \Rightarrow 0 < u < 1$
$\Rightarrow dx = b du$

$$\Rightarrow \Gamma(a) = \int_0^1 u^{a-1} \cdot b^{a-1} e^{-ub} \cdot b du$$

$$= \int_0^1 b^a u^{a-1} e^{-bu} du$$

$\Rightarrow$ Gamma Distr. is normalized

(10) PRML # 2.42, page 134

$$E(\lambda) = \int_0^1 \lambda \frac{1}{\Gamma(a)} b^a \lambda^{a-1} e^{-b\lambda} \, d\lambda$$

$$= \frac{1}{\Gamma(a)} \int_0^1 \lambda \, b^a \lambda^{a-1} e^{-b\lambda} \, d\lambda$$

$$= \frac{1}{\Gamma(a)} \frac{1}{b} \underbrace{\int_0^1 b^{a+1} \lambda^{a} e^{-b\lambda} \, d\lambda}_{①}$$

$$① = \Gamma(a+1)$$

$$\Rightarrow E(\lambda) = \frac{\Gamma(a+1)}{\Gamma(a)} \cdot \frac{1}{b} = \frac{a}{b}$$

$$Var(\lambda) = \int_0^\infty \frac{b^a}{\Gamma(a)} \lambda^{a+1} e^{-b\lambda} - E(x)^2$$

$$= \frac{b^a}{\Gamma(a)} \int_0^\infty \lambda^{a+1} e^{-b\lambda} \, d\lambda - \left(\frac{a}{b}\right)^2$$

let $u = b\lambda \Rightarrow du = b\,d\lambda$

$\Rightarrow b = u/\lambda \quad \lambda = u/b$

$\Rightarrow \mathrm{var}(u)$

$$= \frac{b^a}{\Gamma(a)} \int_0^\infty \left(\frac{u}{b}\right)^{a+1} e^{-u} \, b\,du - \left(\frac{a}{b}\right)^2$$

$$= \frac{b^a}{b^{a+2}\Gamma(a)} \int_0^\infty u^{a+1} e^{-u} \, du - \left(\frac{a}{b}\right)^2$$

$$= \frac{1}{b^2\Gamma(a)} \int_0^\infty u^{a+1} e^{-u} \, du - \left(\frac{a}{b}\right)^2$$

$$= \frac{1}{b^2\Gamma(a)} \Gamma(a+2) - \left(\frac{a}{b}\right)^2$$

$b^2 \Gamma(a)$

$$= \frac{\Gamma(a+2) - a^2\Gamma(a)}{b^2\Gamma(a)}$$

$$= \frac{a(a+1)\Gamma(a) - a^2\Gamma(a)}{b^2\Gamma(a)}$$

$$= \frac{a^2+a-a^2}{b^2} = \boxed{a/b^2 = \mathrm{var}(E)}$$

$$\frac{d}{d\lambda} \mathrm{Gamma}(\lambda \mid a,b) = \frac{d}{d\lambda}\left(b^a \lambda^{a-1} e^{-b\lambda} \right) = 0$$

$$= (a-1)\lambda^{a-2} e^{-b\lambda} - b e^{-b\lambda} \lambda^{a-1} = 0$$

$$= \lambda^{a-2} e^{-b\lambda} (a-1-b\lambda) = 0$$

$$\Rightarrow \boxed{\lambda = \frac{a-1}{b}}$$