

MATH 557 - ASSIGNMENT 2 SOLUTIONS

(a) The likelihood, log-likelihood and its derivatives are

$$L(\lambda|\underline{x}) = \lambda^n \exp \left\{ -\lambda \sum_{i=1}^n x_i \right\}$$

$$l(\lambda|\underline{x}) = n \log \lambda - \lambda \sum_{i=1}^n x_i$$

$$\dot{l}(\lambda|\underline{x}) = n/\lambda - \sum_{i=1}^n x_i$$

$$\ddot{l}(\lambda|\underline{x}) = -n/\lambda^2$$

Solving $\dot{l}(\lambda|\underline{x}) = 0$ yields the ML estimator as

$$\hat{\lambda}_{ML} = \frac{n}{\sum_{i=1}^n X_i} = 1/\bar{X}.$$

For the the method of moments estimator, we equate sample and theoretical first moments:

$$\bar{X} \quad \text{set equal to} \quad \int_0^\infty x f_{X|\lambda}(x|\lambda) dx = 1/\lambda$$

so that

$$\hat{\lambda}_{MM} = 1/\bar{X}.$$

2 MARKS

(b) The Bayes estimate under squared error loss is (from notes) the posterior mean: here

$$\begin{aligned} \pi_{\lambda|\underline{X}}(\lambda|\underline{x}) &\propto L(\lambda|\underline{x})\pi_{\lambda}(\lambda) \\ &= \lambda^n \exp \left\{ -\lambda \sum_{i=1}^n x_i \right\} \alpha \exp\{-\alpha\lambda\} \\ &\propto \lambda^n \exp \left\{ -\lambda(\alpha + \sum_{i=1}^n x_i) \right\} \end{aligned}$$

By inspection, we note that is a Gamma distribution, and if $s(\underline{x}) = \sum_{i=1}^n x_i$

$$\pi_{\lambda|\underline{X}}(\lambda|\underline{x}) \equiv \text{Gamma}(n+1, \alpha + s(\underline{x}))$$

yielding the Bayes estimator from the expectation of the distribution as

$$\hat{\lambda}_B = \frac{n+1}{\alpha + S(\underline{X})}$$

3 MARKS

- (c) We need an estimator that has the correct expectation. Note first that, in this model, from standard distributional results (proved by mgfs say)

$$S(\underline{X}) = \sum_{i=1}^n X_i \sim \text{Gamma}(n, \lambda)$$

and therefore by transformation results, the reciprocal of $S(\underline{X})$ is Inverse Gamma distributed:

$$1/S(\underline{X}) \sim \text{IGamma}(n, \lambda)$$

with expectation $\lambda/(n-1)$. Therefore if we define $U = (n-1)/S(\underline{X})$, then U is unbiased for λ . The variance of the estimator is (using standard results for the Inverse Gamma)

$$(n-1)^2 \text{Var}[1/S(\underline{X})] = \frac{(n-1)^2 \lambda^2}{(n-1)^2(n-2)} = \frac{\lambda^2}{(n-2)}$$

3 MARKS

- (d) $\hat{\lambda}_{ML}$ and $\hat{\lambda}_{MM}$ are inverse Gamma distributed, and by the above argument, the estimators have variance

$$\frac{n^2 \lambda^2}{(n-1)^2(n-2)}$$

As $n \rightarrow \infty$, this variance decreases to zero at rate n . For $\hat{\lambda}_B$:

- as $n \rightarrow \infty$, $\hat{\lambda}_B$ and $\hat{\lambda}_{ML}$ are asymptotically equivalent for any finite α , and the variance goes to zero at rate n .
- If $\alpha = 0$, we have an improper prior, but proper posterior, and the estimator has variance

$$\frac{(n+1)^2 \lambda^2}{(n-1)^2(n-2)}$$

As $n \rightarrow \infty$, this variance decreases to zero at rate n . The estimator is asymptotically equivalent to $\hat{\lambda}_{ML}$.

- For finite n , and $\alpha > 0$, we cannot compute the exact frequentist variance analytically. However, it can be easily studied using simulation.

4 MARKS

The following code in R can be used to study the variance of the ML and Bayesian estimators:

```
lam.bayes<-rep(0,4)
n<-10; lambda<-1; alpha<-c(0,1,2)
X<-rexp(n,lambda)
lam.ml<-1/mean(X) for(l in 1:length(alpha)){
  lam.bayes[l]<-(n+1)/(alpha[l]+sum(X))
}
```

See the plots on page 4 for a study of the variance and MSE of the estimators for n ranging from 1 to 50, for different values of λ . The Bayes estimator is constructed using different choices for prior hyperparameter α .

(e) The prior predictive is

$$\begin{aligned} f_{\underline{X}|\alpha}(\underline{x}|\alpha) &= \int_0^\infty \lambda^n \exp\{-\lambda s(\underline{x})\} \alpha \exp\{-\alpha\lambda\} d\lambda \\ &= \frac{\alpha \Gamma(n+1)}{(\alpha + s(\underline{x}))^{n+1}} \end{aligned}$$

as the integrand is proportional to a Gamma pdf.

3 MARKS

(f) Let $s(\underline{x}^*) = \sum_{i=1}^m x_i^*$. The posterior predictive distribution is dependent on the data only through $s(\underline{x}^*)$.

$$\begin{aligned} f_{\underline{X}^*|\underline{X},\alpha}(\underline{x}^*|\underline{x},\alpha) &= \int_0^\infty \lambda^m \exp\{-\lambda s(\underline{x}^*)\} \frac{(\alpha + s(\underline{x}))^{n+1}}{\Gamma(n+1)} \lambda^n \exp\{-\lambda(\alpha + s(\underline{x}))\} d\lambda \\ &= \frac{(\alpha + s(\underline{x}))^{n+1}}{\Gamma(n+1)} \int_0^\infty \lambda^{n+m} \exp\{-\lambda(\alpha + s(\underline{x}) + s(\underline{x}^*))\} d\lambda \\ &= \frac{(\alpha + s(\underline{x}))^{n+1}}{\Gamma(n+1)} \frac{\Gamma(n+m+1)}{(\alpha + s(\underline{x}) + s(\underline{x}^*))^{n+m+1}} \end{aligned}$$

3 MARKS

To understand the large-sample/asymptotic behaviour, consider the distribution of

$$T(\underline{X}^*) = \overline{X}^* = \frac{S(\underline{X}^*)}{m}$$

We have that, given λ ,

$$\begin{aligned} f_{T^*|\underline{X},\alpha}(t|\underline{x},\alpha) &= \int_0^\infty \frac{(mt)^m}{\Gamma(m)} \lambda^{m-1} \exp\{-mt\lambda\} \frac{(\alpha + s(\underline{x}))^{n+1}}{\Gamma(n+1)} \lambda^n \exp\{-\lambda(\alpha + s(\underline{x}))\} d\lambda \\ &= \frac{(mt)^m}{\Gamma(m)} \frac{(\alpha + s(\underline{x}))^{n+1}}{\Gamma(n+1)} \int_0^\infty \lambda^{n+m-1} \exp\{-\lambda(\alpha + s(\underline{x}) + mt)\} d\lambda \\ &= \frac{(mt)^m}{\Gamma(m)} \frac{(\alpha + s(\underline{x}))^{n+1}}{\Gamma(n+1)} \frac{\Gamma(n+m)}{(\alpha + s(\underline{x}) + mt)^{n+m}} \\ &= \frac{(\alpha + s(\underline{x}))^{n+1}}{\Gamma(n+1)} \frac{\Gamma(n+m)}{\Gamma(m)} \frac{(mt)^m}{(\alpha + s(\underline{x}) + mt)^{n+m}} \\ &\approx \frac{(\alpha + s(\underline{x}))^{n+1}}{\Gamma(n+1)} \frac{m^n}{(\alpha + s(\underline{x}) + mt)^n} \left(1 + \frac{\alpha + s(\underline{x})}{mt}\right)^{-m} \\ &\approx \frac{(\alpha + s(\underline{x}))^{n+1}}{\Gamma(n+1)} t^n \exp\{-t(\alpha + s(\underline{x}))\} \end{aligned}$$

for large m , that is, the same form as the posterior distribution for λ given $\underline{x}$.

2 MARKS

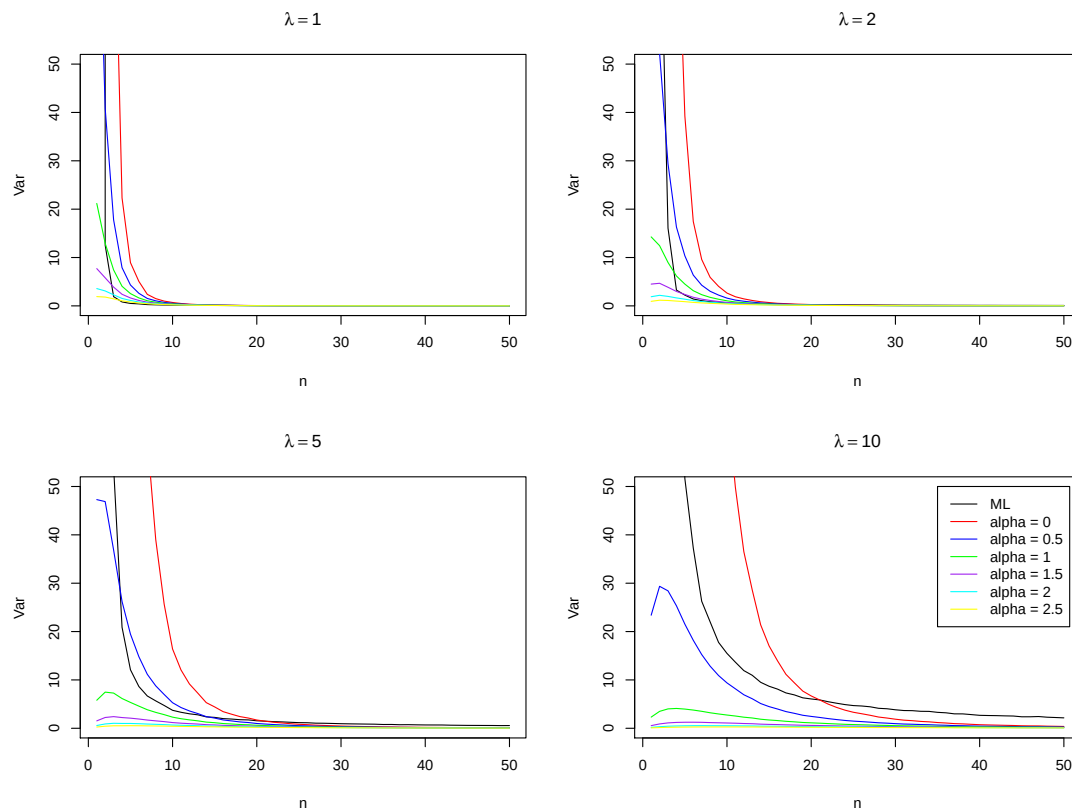


Figure 1: Variance vs n : $\hat{\lambda}_{ML}$ (black) and $\hat{\lambda}_B$ for different choices of α , at four settings of λ .

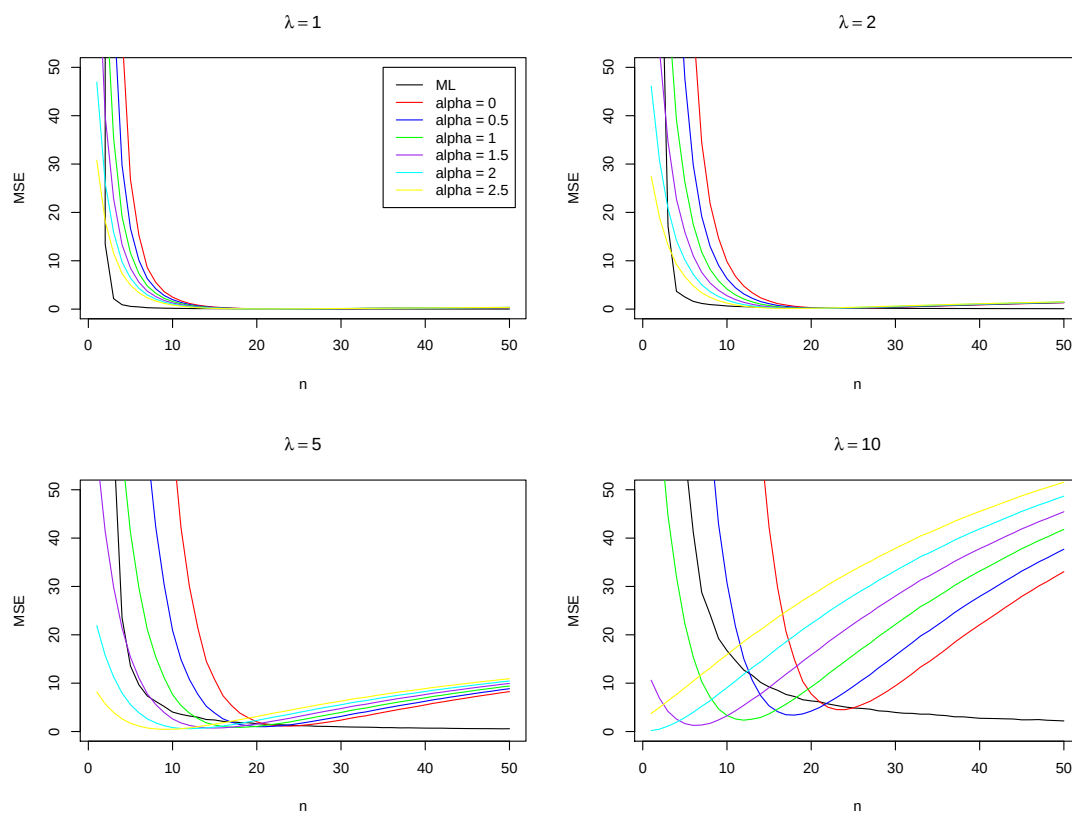


Figure 2: MSE vs n : $\hat{\lambda}_{ML}$ (black) and $\hat{\lambda}_B$ for different choices of α , at four settings of λ .