330:541 Stochastic Signals and Systems Fall 2000

Quiz II and Solutions

You have 2.5 hours to answer the following three questions. You are allowed both sides of an 11x8.5 inch piece of paper for notes. GOOD LUCK!

You may (or may not) find the following theorem useful.

Theorem: If $\hat{\theta} = g(x)$ is a **biased** estimator of θ with $E[\hat{\theta}] = \tau(\theta)$ then

$$\sigma^2_{hat\theta} \geq (\frac{d\tau}{d\theta})^2 \frac{1}{nI}$$

where

$$I = E\left[\left|\frac{\partial L(x,\theta)}{\partial\theta}\right|^2\right]$$

and

$$L(x,\theta) = \log f_{X|\Theta}(x|\theta)$$

1. *(50 points)* **Estimation:** On an observation interval $(0, T)$ you observe arrivals from a Poisson process.

   (a) *(10 points)* Please derive the maximum likelihood estimate of the arrival rate λ. You must show that your answer is unique.

   The conditional distribution is poisson with parameter λT

   $$f_{K|\Lambda}(k|\lambda) = \frac{(\lambda T)^k}{k!} e^{-\lambda T}$$

   We take the first derivative and obtain

   $$\frac{(\lambda T)^k}{k!} e^{-\lambda T}\left[\frac{k}{\lambda} - T\right] = 0$$

   This equation is satisfied at $\lambda = k/T$ and at $\lambda = \infty$. Since the conditional probability $f_{K|\Lambda}(k|\lambda)$ is zero for $\lambda = 0$ and $\lambda = \infty$ and the function must be non-negative, there can exist only one intervening maximum. So the result is unique.

   Note that we could have taken the log of the likelihood function and differentiated that instead. We would find that the resultant *function is actually CONCAVE in λ and thus has a unique maximum. The log likelihood function approach will be illustrated in part c).*

(b) *(10 points)* Show that your estimate is unbiased/biased
$E[\frac{k}{T}] = \frac{\lambda T}{T} = \lambda$ *so unbiased.*

(c) *(10 points)* Show that your estimate is consistent/inconsistent
The problem above was for one observation interval. The consistency problem has to do with multiple independent observations. There are two ways to go here. One can extend the observation interval T directly and show that the estimate is consistent in T. However, we've used a discrete model of consistency, so we'll use n independent observations on disjoint (and therefore independent) intervals of duration T and form $k_1, k_2, \cdots, k_n$.

The joint distribution on the k_i is a product distribution which works out to

$$f_{\mathbf{K}|\Lambda}(\mathbf{k}|\lambda) = \frac{(\lambda T)^{\sum_i k_i}}{\prod_i (k_i)!} e^{-n\lambda T}$$

Taking the log we have

$$L(\mathbf{x}, \lambda) = (\sum_i k_i)(\log \lambda + \log T) - \sum_i \log(k_i)! - n\lambda T$$

Differentiating with respect to λ and setting to zero we obtain

$$\frac{\sum_i k_i}{\lambda} - nT = 0$$

We note that the second derivative is strictly negative so the function is concave and the solution to the extremal equation is unique. In any case, we then have

$$\hat{\lambda} = \frac{\sum_i k_i}{nT}$$

This is unbiased as before since $E[k_i] = \lambda T$ and since the variance of independent random variables is the sum of the variances we have

$$\sigma^2_{\hat{\lambda}|} = \frac{\lambda}{nT}$$

So the estimate is consistent. You'll also note that could rewrite the previous expression as

$$\sigma^2_{\hat{\lambda}|} = \frac{\lambda}{\mathcal{T}}$$

where $\mathcal{T}$ is a longer interval of size nT. That is, the n independent intervals could have been one interval of size nT and the n observations could have been lumped into a single observation of the sum of arrivals in each smaller interval. Thus, the estimate is also consistent in T.

(d) *(20 points)* Show that your estimate is efficient/inefficient

All we have to do is calculate I.

$$nI = E\left[|\frac{\partial L(\mathbf{x},\lambda)}{\partial \lambda}|^2\right] = E\left[|\frac{\sum_i k_i}{\lambda} - nT|^2\right] = \frac{1}{\lambda^2}E\left[(\sum_i k_i - n\lambda T)^2\right]$$

This is just the variance of a Poisson variable $\sum_i k_i$ on an extended interval nT and thus we have $nI = \frac{n\lambda T}{\lambda^2} = nT/\lambda$ so that

$$\frac{1}{nI} = \frac{\lambda}{nT}$$

which is identical to the variance of our ML estimator. So our estimator is efficient.

2. *(50 points)* **More Estimation:** For the Poisson problem above you are now given that the *a priori* distribution on the rate λ is exponential with parameter α.

   (a) *(10 points)* Please derive the MAP estimate of the arrival rate λ.

   We could find the conditional distribution $f_{\Lambda|K}(\lambda|k)$, but we remember that owing to the definition of conditional probability, if we want to maximize this expression in λ, it suffices to maximize $f_{\Lambda,K}(\lambda,k) = f_{\Lambda|K}(\lambda|k)f_{\Lambda}(\lambda)$.

   We'll do everything in general for n observations to save time later.

$$f_{\Lambda,K}(\lambda,\mathbf{k}) = \frac{\alpha(\lambda T)^{\sum_i k_i}}{\prod_i (k_i)!}e^{-n\lambda T - \alpha\lambda}$$

We then take the log and differentiate to find the maximum –

$$\frac{\partial}{\partial \lambda} \log f_{\Lambda,K}(\lambda, \mathbf{k}) = \frac{\sum_i k_i}{\lambda} - (nT + \alpha) = 0$$

So, we have

$$\hat{\lambda} = \frac{\sum_i k_i}{nT + \alpha}$$

We note that the second derivative is strictly negative as before so the maximum is unique.

(b) *(10 points)* Show that your estimate is unbiased/biased

The expectation (based on a given value of λ) is $\frac{n\lambda T}{nT+\alpha}$ is clearly biased (we'll always assume the answer is lower that what we actually observe. This makes intuitive sense since the expoential distribution from which λ is drawn is more heavily weighted toward zero.

(c) *(10 points)* Show that your estimate is consistent/inconsistent

The variance of the estimator (again, given a particular value of λ is $\frac{n\lambda T}{(nT+\alpha)^2}$ which goes to zero in n. So the estimate is consistent.

(d) *(20 points)* Show that your estimate is efficient/inefficient

The log likelihood function does not change – it is still defined in terms of the ML function. So I is the same as before – $I = nT/\lambda$. However, we use the biased form of Cramer-Rao given above.

$$\tau(\lambda) = \frac{n\lambda T}{nT + \alpha}$$

so that

$$\left(\frac{\partial \tau(\lambda)}{\partial \lambda}\right)^2 = \left(\frac{nT}{nT + \alpha}\right)^2$$

So we find that any biased estimator must have variance at least as large as

$$\left(\frac{nT}{nT + \alpha}\right)^2 \frac{\lambda}{nT} = \frac{n\lambda T}{(nT + \alpha)^2}$$

which is identical to the variance of our estimator. So the estimate is again efficient

3. *(50 points)* **Rutgera the Hunter:** With the money she earned as an investment advisor, Rutgera Univera, the world famous Rutgers ECE graduate student decided to purchase a sailboat and take a world tour. Unfortunately, being a novice sailor, she became stranded on a South Pacific island and was forced to live off the land.

   Rutgera quickly realized that her survival was dependent on saving energy, and since hunting expends energy she has now decided to develop decision rules to minimize energy when hunting.

   Specifically, at the start of each day, Rutgera takes up her place behind a blind past which a herd of small grazing animals pass. These animals, called Poissrats, emit high pitched squeals as they walk. What Rutgera has learned is that they emit squeals in a Poisson fashion either at rate λ_1 or at rate $\lambda_2 < \lambda_1$, and that on any given day, they are equally likely to squeal at rate 1 or rate 2.

   When they squeal at the larger rate λ_1 they are *coming from* the water hole. Being full of water, they are slower and easier to catch. You can assume Rutgera always catches a Poissrat if they are waterlogged.

   If they are squealing at rate λ_2, the are *going to* the water hole and are a lot harder to catch. Assume Rutgera never catches a Poissrat when they are not full of water. Thus, it is in Rutgera's interest to have a good policy for when to jump from her blind and start chasing Poissrats.

   Assume Rutgera uses a T second observation interval to count squeals and uses the number of squeals k as her decision statistic. Further, assume Rutgera expends 2 units of energy on a successful hunt and gains 5 units of energy from eating the Poissrat. Likewise assume Rutgera expends 10 units of energy on an unsuccessful hunt. Also assume that Rutgera loses 1 unit if she simply sits behind her blind all day.

   (a) *(30 points)* Please derive a hunting policy for Rutgera which maximizes her energy intake based on her observation of k, the number of squeals she hears in a T second interval.

   Let A be the event that the poissrats are waterlogged and squealing at rate λ_1 and let B be the event that they squeal at rate λ_2. Define $P_{H(h)|\cdot}$ as the probability that Rutgera hunts (H) or does not hunt (h) given event $\cdot$.

The cost function is then

$$2\mathcal{C} = 5P_{H|A} - 2P_{H|A} - P_{h|A} - P_{h|B} - 10P_{H|B} = 3P_{H|A} - P_{h|A} - P_{h|B} - 10P_{H|B}$$

The factor of 2 multiplying the cost is due to the equal probability of A and B (the 1/2 was brought over to the left hand side for convenience).
Our observable is the number of squeals heard in a T second interval. The conditional distributions are then

$$f_{K|A}(k|A) = \frac{(\lambda_1 T)^k}{k!} e^{-\lambda_1 T}$$

$$f_{K|B}(k|B) = \frac{(\lambda_2 T)^k}{k!} e^{-\lambda_2 T}$$

From inspection of the cost function, we hunt only for when

$$4f_{K|A}(k|A) \geq -9f_{K|B}(k|B)$$

However, let me expand a bit for clarity.

$$2\mathcal{C} = 3\sum_{k\in H} f_{K|A}(k|A) - \sum_{k\in h} f_{K|A}(k|A) - \sum_{k\in h} f_{K|B}(k|B) - 10\sum_{k\in H} f_{K|B}(k|B)$$

We may rewrite this as

$$2\mathcal{C} = 3\sum_{k\in H} f_{K|A}(k|A) - (1 - \sum_{k\in H} f_{K|A}(k|A)) - (1 - \sum_{k\in H} f_{K|B}(k|B)) - 10\sum_{k\in H} f_{K|B}(k|B)$$

which simplifies to

$$2\mathcal{C} = \sum_{k\in H} \Big(4f_{K|A}(k|A) - 9f_{K|B}(k|B)\Big) - 2$$

so we only decide to hunt if $4f_{K|A}(k|A) - 9f_{K|B}(k|B) \geq 0$. A formal decision rule:

$$\frac{f_{K|A}(k|A)}{f_{K|B}(k|B)} \underset{\textit{do not hunt}}{\overset{\textit{HUNT!}}{\gtrless}} \frac{9}{4}$$

Doing the substitutions we have

$$\left(\frac{\lambda_1}{\lambda_2}\right)^k e^{-(\lambda_1-\lambda_2)T} \quad \underset{\textit{do not hunt}}{\overset{\textit{HUNT!}}{\gtrless}} \quad \frac{9}{4}$$

We can make it a simple comparison by taking the log to obtain

$$k \quad \underset{\textit{do not hunt}}{\overset{\textit{HUNT!}}{\gtrless}} \quad \left(\frac{\lambda_1}{\lambda_2}\right)^k e^{-(\lambda_1-\lambda_2)T} \frac{2\log(\frac{3}{2}) + (\lambda_1-\lambda_2)T}{\log(\frac{\lambda_1}{\lambda_2})} = \Gamma$$

where Γ is now defined as the threshold.

(b) *(20 points)* Please derive an expression for Rutgera's average energy intake under your policy.

We just lift the expression for cost in the previous part and use the threshold given

$$\mathcal{C} = \frac{1}{2}\sum_{k\geq\Gamma}\Big(4f_{K|A}(k|A) - 9f_{K|B}(k|B)\Big) - 1$$