The University of Sydney
School of Mathematics and Statistics

Computer Quiz Week 12

STAT3023/3923/4023: Statistical Inference — Semester 2, 2021

Lecturers: Neville Weber and Michael Stewart

Prepare a PDF (or HTML or Word) report as usual and submit to the Canvas Assignment by **3pm**.

Estimating a Poisson mean using squared-error loss

Suppose $\mathbf{X} = (X_1, \dots, X_n)$ consists of iid Poisson(θ) random variables with unknown mean $\theta \in \mathbb{R}$. Consider the decision problem of estimating θ with squared-error loss,

$$L(d|\theta) = (d - \theta)^2$$

so that risk of an estimator $d(\mathbf{X})$ is the *mean-squared error*:

$$R(\theta|d) = E_\theta \left\{ [d(\mathbf{X}) - \theta]^2 \right\}.$$

Consider two estimators:

- The maximum likelihood estimator $d_{\text{mle}}(\boldsymbol{X}) = \bar{X}$ (the sample mean).
- The Bayes procedure $d_{\text{conj}}(\boldsymbol{X})$ using a unit-mean exponential prior (note that this is an example of a *conjugate* prior). The product of the prior and the likelihood is

$$w(\theta) f_\theta(\boldsymbol{X}) = e^{-\theta} \prod_{i=1}^{n} \frac{e^{-\theta} \theta^{X_i}}{X_i!} = \text{const. } e^{-\theta(n+1)} \theta^{(T+1)-1},$$

where $T = \sum_{i=1}^{n} X_i = n\bar{X}$ is the sample total. Since (as a function of θ) this is proportional to a gamma density with shape $T + 1$ and rate $n + 1$, the posterior mean is

$$\boxed{d_{\text{conj}}(\boldsymbol{X}) = \frac{\text{shape}}{\text{rate}} = \frac{T+1}{n+1}}$$

We are going to approximate *rescaled risk* via simulation and plot it, that is we are going to

step 1.
- define a vector of θ values;
- for each θ value, we will
  - generate `B` samples of size `n` from that Poisson(θ) distribution, and obtain the value of the estimators $d_{\text{mle}}(\boldsymbol{X}) = \bar{X}$ and $d_{\text{conj}}(\boldsymbol{X})$ for each;
  - obtain the average squared error for each estimator;
- we will then plot `n` times the average squared error against θ for both estimators.

We are going to do this for two different sample sizes: `n=5` and `n=50`.

1. Define a vector of θ values using `th=1:100/10`. Define also `L.th=length(th)`, `n=5` and define empty vectors `nMSE.mle=0` and `nMSE.conj=0`. Finally define `B=10000` as the number of simulation iterations for each θ value (if you computer is very slow you can try making `B` smaller, say `B=1000`).

2. Write a "double loop". At the `j`-th iteration of the outer loop,

- initialise `mle=0`, `conj=0`;
- simulate `B` samples of size `n` from the Poisson distribution with mean `th[j]`, obtaining values of each estimator, saving them in the vectors `mle` and `conj` respectively;

- compute `n` times the average mean-squared error between the estimates in `mle` and `th[j]` and save it in `nMSE.mle[j]`; do the same for `conj` and `nMSE.conj[j]`.

3. Plot `nMSE.conj` (on the vertical axis) against `th` (on the horizontal axis) using a blue line. Add to this same graph a plot of `nMSE.conj` against `th` using a red line. Add an informative heading and legend.

4. **Comment:** By interpreting the Bayes procedure as a weighted average, explain why $d_{\text{conj}}(\boldsymbol{X})$ seems to do better than $d_{\text{mle}}(\boldsymbol{X})$ for some values of θ and worse for others.

5. Repeat the loop in question 2 and the plot in question 3 but with a different sample size: `n=50`.

6. **Comment:** The mle $d_{\text{mle}}(\boldsymbol{X})$ is *unbiased* while the Bayes procedure $d_{\text{conj}}(\boldsymbol{X})$ is *biased*. What asymptotic property is illustrated by the difference between this plot and the one in question 3?