

## ICIC Data Analysis Workshop: First Day Solutions

0. The estimator  $\hat{\mu} = \sum_i w_i x_i$  has expectation value  $\langle \hat{\mu} \rangle = \sum_i w_i \langle x_i \rangle = \mu \sum_i w_i$ , so it is *unbiased* if  $\sum_i w_i = 1$ .

The variance of  $\hat{\mu}$  is computed by noting that, if the variance of  $x$  is  $\sigma^2$ , the variance of  $wx$  is  $w^2\sigma^2$  if  $w$  is just a constant, and that for the sum of independent variables, the variances add, so

$$\sigma_{\hat{\mu}}^2 = \sum_i w_i^2 \sigma^2.$$

We want the minimum variance estimator, so we minimise  $\sigma_{\hat{\mu}}^2$ , but subject to the constraint  $\sum_i w_i = 1$ . Hence we solve

$$\frac{\partial}{\partial w_k} \left( \sum_i w_i^2 \sigma^2 - \lambda \sum_i w_i \right) = 0.$$

The weights are independent, so  $\partial w_i / \partial w_k = \delta_{ik}$ , giving

$$2w_k \sigma^2 - \lambda = 0,$$

or  $w_k = \text{constant}$ . Applying the constraint that they sum to unity,  $w_i = 1/N$ , and

$$\hat{\mu} = \frac{1}{N} \sum_i x_i.$$

Its variance is

$$\sigma_{\hat{\mu}}^2 = \sum_i w_i^2 \sigma^2 = \sigma^2 \sum_i \frac{1}{N^2} = \frac{\sigma^2}{N}.$$

If we have repeated experiments, then the distribution of  $\hat{\mu}$  will have a mean of  $\mu$  and variance  $\sigma^2/N$ , and under weak conditions it is gaussian distributed. Thus it looks very like the Bayesian result, except that what we have calculated here is the distribution of the estimator for repeated experiments, given a value of  $\mu$ , whereas the Bayesian analysis gives the posterior distribution of  $\mu$  given the (single) dataset, which is really what we want to infer. See the discussion of confidence intervals in the lectures for more discussion of the differences.

1. The distribution of flux densities of extragalactic radio sources are distributed as a power-law with slope  $-\alpha$ , say. In a non-evolving Euclidean universe  $\alpha = 3/2$  and departure of  $\alpha$  from the value  $3/2$  is evidence for cosmological evolution of radio sources. This was the most telling argument against the steady-state cosmology in the early 1960s (even though they got the value of  $\alpha$  wrong by quite a long way).<sup>1</sup>

Given observations of radio sources with flux densities  $S$  above a known, fixed measurement limit  $S_0$ , what is the best estimate for  $\alpha$ ?

The model probability distribution for  $S$  is

$$p(S|\alpha)dS = (\alpha - 1)S_0^{\alpha-1}S^{-\alpha}dS$$

where the factor  $\alpha - 1$  in front of the terms arises from the normalization requirement (note the integral is over  $S$ , not  $\alpha$ )

$$\int_{S_0}^{\infty} dS p(S|\alpha) = 1.$$

So the likelihood function  $L$  for  $n$  observed sources is

$$L = \prod_{i=1}^n (\alpha - 1)S_0^{\alpha-1}S_i^{-\alpha}$$

with logarithm

$$\ln L = \sum_{i=1}^n [\ln(\alpha - 1) + (\alpha - 1) \ln S_0 - \alpha \ln S_i].$$

Maximising  $\ln L$  with respect to  $\alpha$ :

$$\frac{\partial}{\partial \alpha} \ln L = \sum_{i=1}^n \left( \frac{1}{\alpha - 1} + \ln S_0 - \ln S_i \right) = 0$$

---

<sup>1</sup>Note that in this question we assume measurement errors are negligible. In more realistic cases, the errors have to be included - the expected observed distribution would be the true distribution convolved with the error distribution, and the observational cuts then applied. It is slightly more involved, but conceptually not very different.

we find the minimum when

$$\alpha = 1 + \frac{n}{\sum_{i=1}^n \ln \frac{S_i}{S_0}}.$$

Suppose we only observe one source with flux twice the cut-off,  $S_1 = 2S_0$ , then

$$\alpha = 1 + \frac{1}{\ln 2} = 2.44$$

but with a large uncertainty. Clearly, as  $S_i = S_0$  we find  $\alpha \rightarrow \infty$  as expected. In fact  $\alpha = 1.8$  for bright radio sources at low frequencies, significantly steeper than 1.5.

To get a rough estimate of the width of the credibility interval for  $\alpha$ , we can compute

$$\sigma_\alpha^{-2} \simeq -\frac{\partial^2 \ln L}{\partial \alpha^2} = \sum_{i=1}^n \frac{1}{(\alpha - 1)^2},$$

evaluated at the maximum likelihood solution. This is not very accurate as the likelihood shape is not gaussian. For the single source,  $\alpha = 2.44$  and  $n = 1$ , so the (rough) estimate of the error is

$$\sigma_\alpha = 1.44.$$

2. Let the doors be labelled  $a, b, c$ , where  $a$  is the door you choose initially, and  $b$  is the door which is opened. Many, if not all, of the probabilities below should be interpreted as ‘given that you have chosen  $a$ ’, but for clarity we won’t write this explicitly.

Let  $p(a)$  = probability that  $a$  leads to the desired prize, etc.

Let  $B$  be the event that door  $b$  gets opened and leads to worthless junk.

What you want is the probability that  $a$  leads to the prize, given that  $b$  is opened and leads to junk. i.e. the aim is to calculate

$$p(a|B).$$

We can use Bayes' theorem for this:

$$p(a|B) = \frac{p(a, B)}{p(B)} = \frac{p(B|a)p(a)}{p(B)}$$

Now, clearly  $p(a) = p(b) = p(c) = 1/3$  (all doors are equally likely, before any experiment is done).

$p(B|a)$  = probability that door  $b$  is opened, given that  $a$  leads to the prize. Evidently

$$p(B|a) = \frac{1}{2} :$$

Alan could have opened either door  $b$  or  $c$ , since they both lead to junk.

What about  $p(B)$ ? It is the sum of all the joint probabilities:

$$p(B) = p(B, a) + p(B, b) + p(B, c) = p(B|a)p(a) + p(B|b)p(b) + p(B|c)p(c),$$

each of which we can calculate.  $p(a) = p(b) = p(c) = 1/3$ , as before, and  $p(B|a) = 1/2$  as before. Now

$$p(B|b) = 0 :$$

Alan will not open  $b$  since it leads to the prize in this case.

$p(B|c)$  is the most interesting. Given that you have chosen  $a$  (remember this is implicit throughout), then if  $c$  leads to the prize, then Prof Heavens *must* open door  $b$ , i.e.

$$p(B|c) = 1$$

So the probability that your original choice  $a$  leads to the prize is

$$\begin{aligned} p(a|B) &= \frac{p(B|a)p(a)}{p(B|a)p(a) + p(B|b)p(b) + p(B|c)p(c)} & (1) \\ &= \frac{\frac{1}{2} \frac{1}{3}}{\frac{1}{2} \frac{1}{3} + \left(0 \times \frac{1}{3}\right) + \left(1 \times \frac{1}{3}\right)} \\ &= \frac{1}{3} \end{aligned}$$

So you would double your chances (from 1/3 to 2/3) if you switch to the other door.

3. Use the product rule:

$$P(M, N) = P(M)P(N|M) \quad (2)$$

and the first term is a Poisson distribution (given), with parameter (=mean)  $\mu = \lambda t$ . The second term is just a binomial - the probability of detecting  $N$  photons from  $M$ , if the individual probability is  $p$ .

Thus

$$p(M, N) = \frac{\mu^M}{M!} e^{-\mu} \frac{M!}{N!(M-N)!} p^N q^{M-N} \quad (3)$$

To get  $P(N)$ , marginalise over  $M$ , noting that  $M$  is *discrete*, so it is a sum, not an integral:

$$P(N) = \sum_{M=N}^{\infty} P(M, N) \quad (4)$$

where we note that at least  $M$  photons must be emitted if  $N$  are detected... Substituting for  $P(M, N)$ , and letting  $i = M - N$ ,

$$P(N) = \sum_{M=N}^{\infty} p(M, N) = \frac{p^N q^{-N} e^{-\mu} (\mu q)^N}{N!} \sum_{i=0}^{\infty} \frac{(\mu q)^i}{i!}. \quad (5)$$

The sum is just  $e^{\mu q}$ , hence

$$P(N) = \frac{(\mu p)^N e^{-\mu p}}{N!}, \quad (6)$$

which is just a Poisson distribution with mean  $\mu p$ .

The last part follows similar lines, and is left for you to work out.