APTS Statistical Inference, assessment questions, 2014

1. Here is a question on credible sets and transformation invariance.

   (a) Let $\boldsymbol{X} := (X_1, \dots, X_m)$. Suppose you believe $\boldsymbol{X} \sim f_{\boldsymbol{X}}$ for some specified PMF $f_{\boldsymbol{X}}$. Define a level-β *credible set* for $\boldsymbol{X}$.

   (b) In the special case where $m = 1$, define a level-β *equitailed credible interval* for X_1. Explain what it would mean for a credible interval to be *transformation-invariant*, and show that equitailed credible intervals are transformation-invariant. Hint: think about a bijective tranformation $g : x \mapsto y$.

   (c) In the general case ($m \geq 1$) a level-β *high probability credible set* is defined as

   $$\mathcal{C}_\beta := \big\{\boldsymbol{x} \in \mathcal{X} : f_{\boldsymbol{X}}(\boldsymbol{x}) \geq c\big\}$$

   for the largest c satisfying $\Pr(\boldsymbol{X} \in \mathcal{C}_\beta) \geq \beta$.

      i. Sketch this credible set in the case where $m = 1$.

      ii. Show that it is the smallest set which contains at least β of the total probability.

      iii. Show that it is not transformation-invariant.

   ('Show' can be informal, but it should still be compelling.)

   (d) Suppose you only know $f_{\boldsymbol{X}}$ up to a multiplicative constant, i.e.

   $$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{1}{z}\, g(\boldsymbol{x})$$

   where you can evaluate g, and, necessarily, $z = \sum_{\boldsymbol{x}} g(\boldsymbol{x})$, but it may be very expensive to evaluate z. Using g, you have constructed an MC sampler (possibly MCMC) for $f_{\boldsymbol{X}}$. How would you use the output from this sampler to identify that subset of $\mathcal{X}$ which was in the 95% high probability credible set?

   (e) Your interest is in a 95% credible set for X_1. Using your MC sampler, you have a 95% high probability credible set for $\boldsymbol{X}$, say $\mathcal{C}_{0.95}$, which you can then marginalise, to get

   $$\mathcal{C}^a_{0.95} := \big\{x_1 : \boldsymbol{x} \in \mathcal{C}_{0.95}\big\}.$$

But you can also create a 95% high probability credible set directly from the sampled x_1 values alone, by estimating f_{X_1} from the sample (e.g. using a histogram estimator, or a kernel density estimator). Call this credible set $\mathcal{C}^b_{0.95}$. Which of these two credible sets for X_1 do you expect to be smaller?

(f) (Optional) Design and execute a simulation experiment to test your answer to the previous question.

(g) Suppose now that you have a parametric model $f_{\mathbf{X}}(\cdot;\theta)$ for $\theta \in \Omega$, where Ω is some convex subset of $\mathbb{R}^p$. The likelihood function is $L(t) := \Pr(Q;t)$ where Q is the data proposition. Show that the level sets of L are transformation-invariant.

(h) Use L.J. Savage's stable estimation theorem (Edwards *et al.*, 1963) to explain why it is that when the likelihood is concentrated into a small enough region of the parameter space, the high posterior density credible sets are nearly transformation-invariant.

2. This is a question about Lindley's paradox (Lindley, 1957), P-values, and the conventional 0.05 threshold.

(a) Suppose you are sitting in a bar talking to an experimental psychologist about significance levels. An informal statement of *Lindley's paradox* is that a P-value for H_0 smaller than α can correspond to a likelihood ratio for H_0 versus H_1 greater than $1/\alpha$. Provide an proof of this statement which you can sketch on a napkin (by all means include the napkin in your answer.) Hint: see DeGroot and Schervish (2002, sec. 8.9), or work backwards from the next question.

(b) Study Figure 1, and see if you can replicate it, either showing your workings or including your code. Some hints:

$$\bar{X} \sim \mathrm{N}(\mu, \sigma^2/n) \tag{1a}$$

with σ^2 known (take $\sigma = 1$). The competing hypotheses are

$$H_0 : \mu = 0 \quad \text{versus} \quad H_1 : \mu = 1 \tag{1b}$$

(i.e. a separation of σ). What is this graph showing?

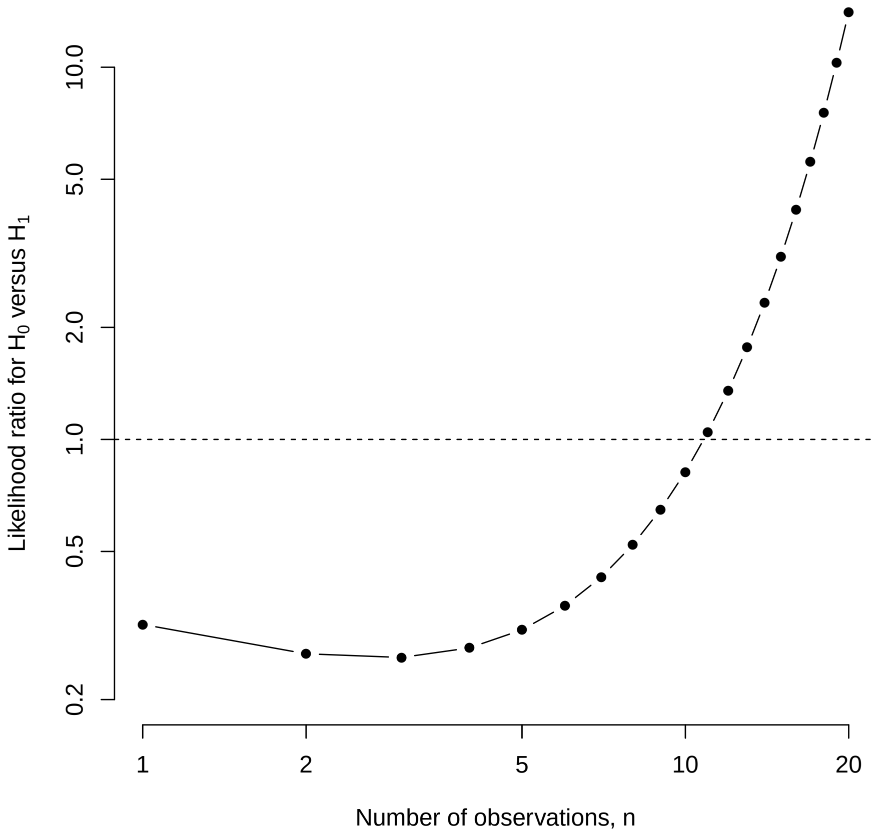

Figure 1: The likelihood ratio corresponding to a P-value for H_0 of 0.05, for the model and hypotheses given in (1).

(c) Consider

$$H_0 : \mu = 0 \quad \text{versus} \quad H_1 : \mu > 0$$

in the case where σ^2 is known and n is fixed. Produce a graph showing the value of the minimum likelihood ratio over H_1 for a range of P-values for H_0 from 0.001 to 0.1. Check your graph against the minimum shown in Figure 1. Hint: you should be able to compute this graph directly. You might find Edwards *et al.* (1963) or Goodman (1999a,b) helpful.

Comment on whether the conventional choice of 0.05 is a suitable threshold for choosing between hypotheses, or whether some other choice might be better. You may also like to reflect on the origin of the value 0.05, see Cowles and Davis (1982).

(d) (Optional) Explain the 'calibration' approach of Sellke *et al.* (2001, sec. 3.1) for providing a lower bound on a likelihood ratio based on a P-value for a simple hypothesis.

Jonathan Rougier
University of Bristol
Dec 2014

References

M.S. Bartlett, 1957. A comment on D.V. Lindley's statistical paradox. *Biometrika*, **44**, 533–534.

M. Cowles and C. Davis, 1982. On the origins of the .05 level of statistical significance. *American Psychologist*, **37**(5), 553–558.

M.H. DeGroot and M.J. Schervish, 2002. *Probability and Statistics*. Addison-Wesley Publishing Co., Reading MA, 3rd edition.

W. Edwards, H. Lindman, and L.J. Savage, 1963. Bayesian statistical inference for psychological research. *Psychological Review*, **70**(3), 193–242.

S. Goodman, 1999a. Toward evidence-based medical statistics. 1: The p-value fallacy. *Annals of Internal Medicine*, **130**, 995–1004.

S. Goodman, 1999b. Toward evidence-based medical statistics. 2: The Bayes factor. *Annals of Internal Medicine*, **130**, 1005–1013.

D.V. Lindley, 1957. A statistical paradox. *Biometrika*, **44**, 187–192. See also Bartlett (1957).

T. Sellke, M.J. Bayarri, and J.O. Berger, 2001. Calibration of p values for testing precise null hypotheses. *The American Statistician*, **55**(1), 62–71.