Theory of Inference: Homework 5

Here are two exam-style revision questions, about p-values and confidence sets.

1. (a) Consider the general model in which $(X_1, \dots, X_m) \sim \mathrm{p}(\boldsymbol{x}; \theta)$ for $\theta \in \Omega$, with observables $Y_i := g_i(\boldsymbol{X})$ for $i = 1, \dots, n$, where the g_i are specified functions of $\boldsymbol{x}$. State the general formula for $\mathrm{p}(\boldsymbol{y}; \theta)$, and also the special case where $\boldsymbol{X} \overset{\text{iid}}{\sim} \mathrm{p}(x; \theta)$ and $Y_i = X_i$ for $i = 1, \dots, n$. [5 marks]

   (b) (i) Consider the model $\boldsymbol{Y} \sim \mathrm{p}(\boldsymbol{y}; \theta)$ for $\theta \in \Omega$. Under what conditions is the statistic $p_0(\boldsymbol{y})$ a P-value for the simple hypothesis $H_0 : \theta = \theta_0$?

   (ii) Let $t(\boldsymbol{y})$ be any statistic. Prove that

$$p_0(\boldsymbol{y}) := \Pr\left\{t(\boldsymbol{Y}) \geq t(\boldsymbol{y}); \theta_0\right\}$$

   is a P-value for H_0. You may take as given the Probability Integral Transform (PIT), which states that if F_X is the distribution function of X, then $F_X(X)$ has a sub-uniform distribution.

   (iii) Give an example of a P-value which is completely uninformative about H_0, and explain how this possibility affects our interpretation of P-values.

   [10 marks]

   (c) Let $p(\boldsymbol{y}; \theta_0)$ be a P-value for $H_0 : \theta = \theta_0$, and suppose that this can be computed for each $\theta_0 \in \Omega_0 \subset \Omega$. Define what is meant by a P-value for $H_0 : \theta \in \Omega_0$, and show that

$$p_{\Omega_0}(\boldsymbol{y}) := \sup_{\theta_0 \in \Omega_0} p(\boldsymbol{y}; \theta_0)$$

   is such a P-value. [5 marks]

   (d) You have computed $p_0(\boldsymbol{y}^{\text{obs}}) = 0.0135$ for some dataset $\boldsymbol{y}^{\text{obs}}$. Interpret this value for your non-statistical client in the case where p_0 is an exact P-value for H_0, and the case where p_0 is not an exact P-value. [5 marks]

2. Consider a statistical model of the form $\boldsymbol{Y} \sim \mathrm{p}(\cdot; \theta)$ for $\theta \in \Omega \subset \mathbb{R}^p$.

   (a) Let $\mathcal{C}$ be a function mapping from $\boldsymbol{\mathcal{Y}}$ to subsets of Ω. Define the *coverage* of $\mathcal{C}$ at θ. Define a level-β *confidence set* for θ. What special property does an *exact* confidence set have? [6 marks]

   (b) Propose an exact 95% confidence set for θ which is nonetheless entirely uninformative about θ. What do you conclude from the fact that this is possible? [6 marks]

   (c) State and prove the *marginalisation theorem* for confidence sets. [6 marks]

   (d) Describe a general-purpose approach for computing a 95% confidence set for θ, based on level sets of the form

$$\mathcal{C}(\boldsymbol{y}) := \left\{ \theta : \log \mathrm{p}(\boldsymbol{y}; \theta) \geq \log \mathrm{p}(\boldsymbol{y}; \hat{\theta}(\boldsymbol{y})) - k \right\},$$

   where $\hat{\theta}(\boldsymbol{y})$ is the Maximum Likelihood (ML) estimator for θ. Include in your description a justification for the form given above, an explanation of *level error*, and a sampling-based approach for reducing level error. [7 marks]

If you would like to hand in this homework for marking, please do so by 5pm on Wed 6 May, in the box outside my offce.

Jonathan Rougier
Mar 2015