APTS Statistical Inference, Preliminary Questions

Jonathan Rougier, 2012.

These questions are to prepare you for the APTS module Statistical Inference. You should attempt them after reading sections 1 and 2.1 of the notes.

1. *Basic probability theory.*

   You can look up the answers to these questions in a standard textbook, but it might be fun to try them first.

   (a) Markov's inequality states that if X is a non-negative random variable with finite mean μ, then $\Pr\{X \geq a\} \leq \mu/a$, for all $a > 0$. Prove this.

   (b) Now prove Chebyshev's inequality using Markov's inequality.

   (c) Now prove that if $X_1, X_2, \dots$ are an IID sequence with finite variance, then

   $$n^{-1}(X_1 + \cdots + X_n) \xrightarrow{\text{P}} \mu,$$

   the Weak Law of Large Numbers (WLLN), where μ is the mean of the X's.

   (d) X_n converges to Y in quadratic mean, written $X \xrightarrow{\text{qm}} Y$, if

   $$E\big\{(X_n - Y)^2\big\} \to 0.$$

   Use Chebyshev's inequality to prove that $X_n \xrightarrow{\text{qm}} Y$ implies $X_n \xrightarrow{\text{P}} Y$.

   (e) The mean squared error (MSE) of an estimator s is defined as (I'm dropping the '$;\theta$' argument, for clarity)

   $$\text{MSE}(S) := E\big\{(S - \theta)^2\big\}.$$

   Show that $\text{MSE}(S) = \text{bias}(S)^2 + \text{Var}(S)$.

   (f) Show that if the bias and the standard error of an estimator both go to zero as n increases for all θ, then the estimator is consistent; i.e. $S \xrightarrow{\text{P}} \theta$ for all $\theta \in \Omega$.

---

[0]Comments on this document are welcome; please address them by email to me at `j.c.rougier@bristol.ac.uk`. This version of the document created on November 27, 2012.

(g) The formal statement of Jensen's inequality is that if g is a convex function and $(p_1, \ldots, p_m)$ is a probability assignment over $\mathfrak{X}$, then

$$E\{g(\boldsymbol{X})\} = \sum_{j=1}^{m} p_j g(\boldsymbol{x}_j) \geq g\Big(\sum_{j=1}^{m} p_i\, \boldsymbol{x}_j\Big) = g(E\{\boldsymbol{X}\}).$$

Prove this. (Hint: start with the definition of a convex function for which $m = 2$, and then use induction.)

(h) Now prove Gibbs's Inequality.

2. *Fun with the Poisson distribution.*

The Poisson distribution has

$$f_X(x; \lambda) = \exp(-\lambda)\frac{\lambda^x}{x!} \quad \text{for } x = 0, 1, 2, \ldots,$$

and zero otherwise, where $\lambda > 0$.

(a) Show that the moment generating function (MGF) of the Poisson distribution is

$$M(t; \lambda) = \exp\{\lambda(e^t - 1)\}.$$

And hence show that $E(X; \lambda) = \mathrm{Var}(X; \lambda) = \lambda$.

(b) Suppose that $\boldsymbol{X} \overset{\text{iid}}{\sim} f_X(x; \lambda)$ and that $Y_n := X_1 + \cdots + X_n$. Prove that $Y \sim f_X(y; n\lambda)$ and that $Y_n \xrightarrow{\text{D}} \mathrm{N}(y; n\lambda, n\lambda)$.

(c) Find the score function for $X \sim f_X(x; \lambda)$. Confirm that its expectation is zero. Show that its variance is $1/\lambda =: i_1(\lambda)$, where i_1 is the Fisher Information of X.

(d) Show that if $\boldsymbol{X} \overset{\text{iid}}{\sim} f_X(x; \lambda)$, then the MLE for λ is $\hat{\lambda}(\boldsymbol{x}) = \bar{x}$, where $\bar{x}$ is the sample mean.

(e) Show that $\hat{\lambda}(\boldsymbol{x})$ is unbiased, and that the standard deviation of $\hat{\lambda}(\boldsymbol{x})$ is $\sqrt{\lambda/n}$. Confirm that $\hat{\lambda}(\boldsymbol{x})$ achieves the Cramér-Rao lower bound.

(f) Compute the expression $u(\boldsymbol{x}, \lambda)^2/i_n(\lambda)$. Identify its asymptotic distribution.