Theory of Inference: Homework 1

Here are some questions about some of the most important inequalities in the whole of probability and statistics. Before you do these questions, make sure you have looked at the self-assessment questions on chs 1 to 3.

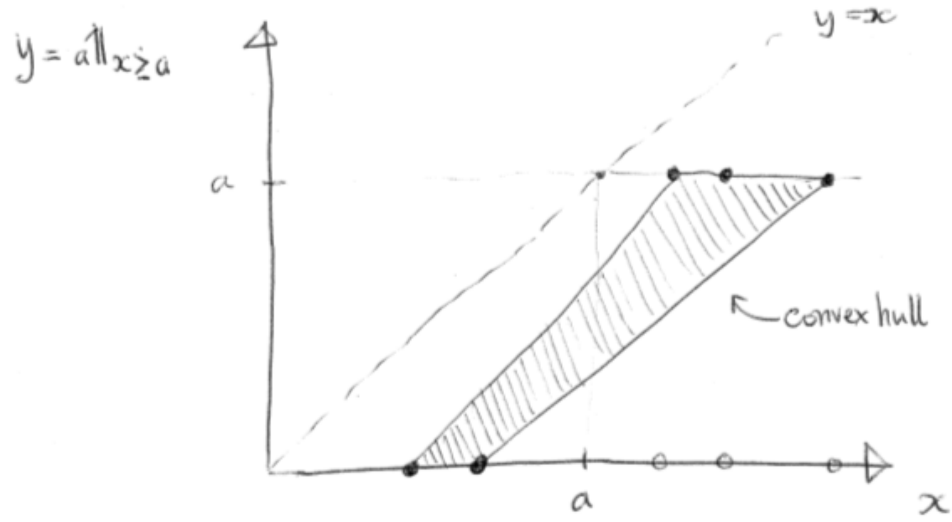
1. Give a complete statement of the Fundamental Theorem of Prevision (FTP) in its 'convex hull' form.

   Answer. Let $\boldsymbol{X} := (X_1, \dots, X_m)$ be a set of random quantities with joint realm
   $$\boldsymbol{\mathcal{X}} := \left\{\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(s)}\right\} \qquad \boldsymbol{x}^{(j)} \in \mathbb{R}^m.$$
   Let $Z_1 := g_1(\boldsymbol{X}), \dots, Z_k := g_k(\boldsymbol{X})$ be k random quantities, for specified functions $g_1, \dots, g_k$. Define the $(k \times s)$ matrix G with $G_{ij} \leftarrow g_i(\boldsymbol{x}^{(j)})$. The 'original' FTP states that $\boldsymbol{v} \in \mathbb{R}^k$ is a coherent expectation for $(Z_1, \dots, Z_k)$ if and only if
   $$G\boldsymbol{p} = \boldsymbol{v} \qquad \text{for some } \boldsymbol{p} \in \mathbb{S}^{s-1},$$
   where $\mathbb{S}^{s-1}$ is the $(s-1)$-dimensional unit simplex,
   $$\mathbb{S}^{s-1} := \left\{\boldsymbol{p} \in \mathbb{R}^s : p_j \geq 0, \sum\nolimits_j p_j = 1\right\}.$$
   Since
   $$G\boldsymbol{p} = G_{(1)} \cdot p_1 + \cdots + G_{(s)} \cdot p_s$$
   where $G_{(j)}$ is the jth column of G, the FTP equivalently asserts that $\boldsymbol{v} \in \mathbb{R}^k$ is a coherent expectation for $(Z_1, \dots, Z_k)$ if and only if it lies in the convex hull of the columns of G.

   (The convex hull of a set of points $\mathcal{G}$ with $g_i \in \mathbb{R}^k$ is precisely the set of all points in $\mathbb{R}^k$ that can be represented as a convex combination of the elements of $\mathcal{G}$. Equivalently, it is the smallest convex set of $\mathbb{R}^k$ that contains every point in $\mathcal{G}$.)

2. Study the proof of Jensen's inequality in ch2 of the notes. When you are sure you understand it, use the same approach to prove Markov's inequality (a diagram is required).

Answer. Study my diagram:



We have one non-negative random quantity X, and then take $g_1(x) \leftarrow x$ and $g_2(x) \leftarrow a\mathbb{1}_{x \dot{\geq} a}$, so $k = 2$. The dots show the realm of $(X, a\mathbb{1}_{X \dot{\geq} a})$, i.e. the columns of G in the notation of the previous answer, and the shaded region is the convex hull of the columns of G. This convex hull lies everywhere below the line $y = x$. Hence if (v_1, v_2) is in the convex hull then $v_1 \geq v_2$. Since (v_1, v_2) is a coherent expectation for $(X, a\mathbb{1}_{X \dot{\geq} a})$ if and only if (v_1, v_2) lies in the convex hull, it follows that $\mathrm{E}(X) \geq \mathrm{E}(a\mathbb{1}_{X \dot{\geq} a})$, from which Markov's inequality follows directly. Note that we could achieve an equality by putting a equal to an element of the realm of X.

3. Prove Chernoff's inequality

$$\Pr(X \dot{\geq} a) \leq \inf_{t>0} e^{-at} M_X(t)$$

where M_X is the moment generating function (MGF) of X.

Answer. From Markov's inequality we get the generalised Markov's inequality that

$$\Pr(X \dot{\geq} a) \leq \frac{\mathrm{E}\{g(X)\}}{g(a)}$$

whenever g is non-negative and strictly increasing. So $g(x;t) \leftarrow \exp(tx)$ fits the bill for $t > 0$, and we have

$$\Pr(X \dot{\geq} a) \leq \frac{\mathrm{E}\{\exp(tX)\}}{\exp(ta)} = e^{-at} M_X(t) \quad \text{for all } t > 0,$$

where M_X is the MGF of X. Because this inequality holds for all $t > 0$, it can be minimised over $t > 0$.

4. Use Chernoff's inequality to provide an upper bound for the tail probability of a Poisson distribution with rate λ. Verify this bound using R.

Answer. The MGF of a Poisson with rate λ is $M_X(t) = \exp\{\lambda(e^t - 1)\}$, so

$$e^{-at} M_X(t) = \exp\{\lambda(e^t - 1) - at\}.$$

The necessary condition for minimisation is

$$\frac{\mathrm{d}}{\mathrm{d}t} \lambda(e^t - 1) - at = 0$$

or $t = \log(a/\lambda)$. The sufficient condition confirms that this is a global minimum. For $t > 0$ we must have $a > \lambda$. Hence, on plugging in and rearranging,

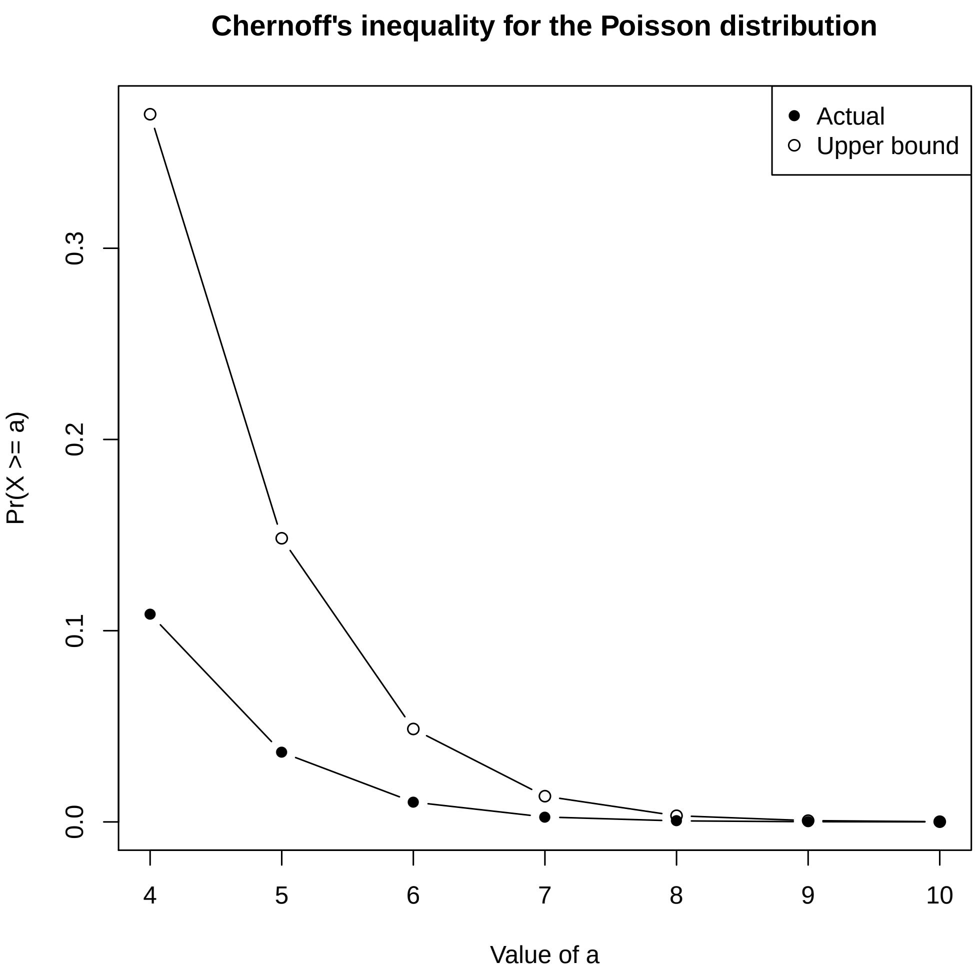
$$\Pr(X \dot{\geq} a; \lambda) \leq e^{a-\lambda} \left(\frac{\lambda}{a} \right)^a \qquad \text{for } a > \lambda.$$

Here is some code to check the answer. Play around with the values. There is a picture after the code.

```
lambda <- 1.8
avals <- 4:10
y <- ppois(avals - 1, lambda, lower.tail = FALSE) # Pr(X >= a)
ub <- exp(avals - lambda) * (lambda / avals)^avals
plot(avals, y, type = "b", pch = 16, ylim = range(c(y, ub)),
  xlab = "Value of a", ylab = "Pr(X >= a)",
  main = "Chernoff's inequality for the Poisson distribution")
points(avals, ub, type = "b", pch = 1)
```

```
legend("topright", legend = c("Actual", "Upper bound"), pch = c(16, 1))

# dev.print(pdf, file = "chernoff.pdf")
```



Please hand in your answers for marking next Tue (24 Feb), at the lecture or by 5pm in the box outside my office door. I will return them in Thu's lecture.

Jonathan Rougier
Feb 2015