Statistics 2 - Problem Sheet 2

Dom Hutchinson

October 15, 2019

Question - 1.

Derive the Maximum Likelihood estimates for the following distributions.

Question 1.1 - $\mathbf{X} \overset{\text{iid}}{\sim} \text{Poisson}(\lambda)$ with λ unknown.

Answer 1.1
Let $\mathbf{X} \overset{\text{iid}}{\sim} \text{Poisson}(\lambda)$ with λ unknown. Then

$$\begin{array}{rrcl} & \ell(\lambda;\mathbf{x}) & = & -\lambda n + \left(\sum_{i=1}^n x_i\right)\ln\lambda + c \\ \implies & \ell'(\lambda;\mathbf{x}) & = & -n + \frac{1}{\lambda}\sum_{i=1}^n x_i \\ \text{Setting} & 0 & = & \ell'(\lambda;\mathbf{x}) \\ \implies & 0 & = & -n + \frac{1}{\hat{\lambda}}\sum_{i=1}^n x_i \\ \implies & \hat{\lambda} & = & \frac{1}{n}\sum_{i=1}^n x_i \\ & & = & \bar{x} \\ \text{We have} & \ell''(\lambda;\mathbf{x}) & = & -\frac{1}{\lambda^2}\sum_{i=1}^n x_i \\ & & < & 0\ \forall\ \lambda > 0 \end{array}$$

Meaning $\hat{\lambda} = \bar{x}$ is a local maximum and thus a maximum likelihood estimate for λ.

Question 1.2 - $X \sim \text{Binomial}(n,p)$ with n known by p unknown.

Answer 1.2
Let $X \sim \text{Binomial}(n,p)$ with n known, but p unknown. Then

$$\begin{array}{rrcl} & \ell(p;x,n) & = & \ln\binom{n}{x} + x\ln p + (n-x)\ln(1-p) + c \\ \implies & \ell'(p;x,n) & = & \frac{n}{p} - \frac{n-x}{1-p} \\ \text{Setting} & 0 & = & \ell'(p;x,n) \\ \implies & 0 & = & \frac{n}{\hat{p}} - \frac{n-x}{1-\hat{p}} \\ \implies & x(1-\hat{p}) & = & \hat{p}(n-x) \\ \implies & x - x\hat{p} & = & \hat{p}n - \hat{p}x \\ \implies & \hat{p} & = & \frac{x}{n} \\ \text{We have} & \ell'(\hat{p};x,b) & = & -\frac{x}{\hat{p}^2} - \frac{n-x}{(1-\hat{p})^2} \\ & & = & -\frac{xn^2}{x^2} - \frac{n-x}{\left(\frac{n-x}{n}\right)^2} \\ & & = & -\frac{n^2}{x} - \frac{n^2}{n-x} \end{array}$$

Since $n \geq x \geq 0$ then $\ell''(\hat{p}; x, n) < 0\ \forall\ n, x$.
Meaning $\hat{p} = \frac{x}{n}$ is a maximum and thus a maximum likelihood estimate for p.

Question 1.3 - $\mathbf{X} \overset{\text{iid}}{\sim} \text{Normal}(\mu, \sigma^2)$ with μ & σ unknown.

Answer 1.3
Let $\mathbf{X} \overset{\text{iid}}{\sim} \text{Normal}(\mu, \sigma^2)$ with μ & σ^2 unknown. Then

$$\begin{array}{rrcl}
 & \ell(\mu, \sigma^2; \mathbf{x}) & = & n \ln \sigma^2 + \frac{1}{\sigma^2} \sum\limits_{i=1}^n (x_i - \mu)^2 + c \\
\implies & \frac{\partial}{\partial \sigma^2} \ell(\mu, \sigma^2; \mathbf{x}) & = & \frac{n}{\sigma^2} - \frac{1}{(\sigma^2)^2} \sum\limits_{i=1}^n (x_i - \mu)^2 \\
\text{Setting} & 0 & = & \frac{\partial}{\partial \sigma^2} \ell(\hat{\mu}, \hat{\sigma}^2; \mathbf{x}) \\
\iff & 0 & = & \frac{n}{\hat{\sigma}^2} - \frac{1}{(\hat{\sigma}^2)^2} \sum\limits_{i=1}^n (x_i - \hat{\mu})^2 \\
\implies & n\hat{\sigma}^2 & = & \sum\limits_{i=1}^n (x_i - \hat{\mu})^2 \\
\implies & \hat{\sigma^2} & = & \frac{1}{n} \sum\limits_{i=1}^n (x_i - \hat{\mu})^2 \\
\text{We have} & \frac{\partial}{\partial \mu} \ell(\mu, \sigma^2; \mathbf{x}) & = & -\frac{2}{\sigma^2} \sum\limits_{i=1}^n (x_i - \mu) \\
\text{Setting} & 0 & = & \frac{\partial}{\partial \mu} \ell(\hat{\mu}, \hat{\sigma}^2; \mathbf{x}) \\
\iff & 0 & = & -\frac{2}{\hat{\sigma}^2} \sum\limits_{i=1}^n (x_i - \hat{\mu}) \\
\implies & 0 & = & \sum\limits_{i=1}^n (x_i - \hat{\mu}) \\
\implies & 0 & = & \left(\sum\limits_{i=1}^n x_i \right) - n\hat{\mu} \\
\implies & \hat{\mu} & = & \frac{1}{n} \sum\limits_{i=1}^n x_i
\end{array}$$

Question 1.4 - $\mathbf{X} \overset{\text{iid}}{\sim} \text{Uniform}(a, b)$ with a & b unknown.

Answer 1.4
Let $\mathbf{X} \overset{\text{iid}}{\sim} \text{Uniform}[a, b]$ with a & b unknown, and $a \leq b$.
We have $L(a, b; \mathbf{x}) = \dfrac{1}{(b-a)^n}$ if $\forall x_i \in \mathbf{x},\ x_i \in [a, b]$.
To maximise $L(a, b; \mathbf{x})$ we want to minimise $(b - a)$.
We should not the further constraints that $a \leq \min\{x_i : i \in [1, n]\}$ & $b \geq \max\{x_i : i \in [1, n]\}$ in order for the sample $x_1, \dots, x_n$ to be valid.
For fixed b we want to maximise $a \implies \hat{a} = \min\{x_i : i \in [1, n]\}$.
For fixed a we want to minimise $b \implies \hat{b} = \max\{x_i : i \in [1, n]\}$.

Question - 2.

Let $X \sim \text{Pareto}(x_0, \theta)$ where $x_0 > 0$ and $\theta > 0$.
Note that

$$f_X(x; x_0; \theta) = \frac{\theta x_0^\theta}{x^{\theta+1}} \mathbb{1}(x \geq x_0)$$

Question 2.1 - Show that the cumulative density function of X is

$$F_X(x; x_0, \theta) = \left\{ 1 - \left(\frac{x}{x_0} \right)^\theta \right\} \mathbb{1}(x \geq x_0)$$

Answer 2.1

$$\begin{aligned}
F_X(x;x_0,\theta) &= \int_{-\infty}^{x} f_X(t;x_0,\theta)dt \\
&= \int_{-\infty}^{x} \frac{\theta x_0^\theta}{t^{\theta+1}}\mathbb{1}(t\geq x_0)dt \\
&= \theta x_0^\theta \int_{-\infty}^{x} t^{-(\theta+1)}\mathbb{1}(t\geq x_0)dt \\
&= \theta x_0^\theta \left[\frac{t^{-(\theta+1)+1}}{-(\theta+1)+1}\right]_{x_0}^{x} \\
&= \theta x_0^\theta \left[\frac{t^{-\theta}}{\theta}\right]_{x_0}^{x} \\
&= \theta x_0^\theta \left[\frac{1}{-\theta x^\theta} - \frac{1}{-\theta x_0^\theta}\right] \\
&= \theta x_0^\theta \left[\frac{1}{\theta x_0^\theta} - \frac{1}{\theta x^\theta}\right] \\
&= 1-\left(\frac{x_0}{x}\right)^\theta \mathbb{1}(x\geq x_0)
\end{aligned}$$

Question 2.2 - Show that the quantile function for X is

$$F_X^{-1}(u;x_0,\theta) = x_0(1-u)^{-\frac{1}{\theta}}$$

Answer 2.2

$$\begin{aligned}
\text{Set} \quad F_X(x;x_0,\theta) &= u \\
\implies \quad 1=\left(\frac{x_0}{x}\right)^\theta &= u \\
\implies \quad \frac{x_0}{x} &= (1-u)^{\frac{1}{\theta}} \\
\implies \quad x &= x_0(1-u)^{-\frac{1}{\theta}} \\
\implies \quad F_X^{-1}(u;x_0,\theta) &= x_0(1-u)^{-\frac{1}{\theta}}
\end{aligned}$$

Question 2.3 - How can we generate random quantities from this distribution?
Hint - Show that the cumulative density function of random variable $Y := F_X^{-1}(U)$ for $U \sim \text{Uniform}[0,1]$ is the same as $X \sim \text{Pareto}(x_0,\theta)$.

Answer 2.3
Let $U \sim \text{Uniform}[0,1]$, $Y := F_X^{-1}(U)$ and $y \in [x_0,\infty)$ be a realisation of Y. Then

$$\begin{aligned}
F_Y(y) &= \mathbb{P}(Y\leq y) \\
&= \mathbb{P}(F_X^{-1}(U)\leq y) \\
&= \mathbb{P}(U\leq F_X(y)) \\
&= \begin{cases} 0 & \text{, if } F_X(y)<0 \\ F_X(y) & \text{, if } F_X(y)\in[0,1] \\ 1 & \text{otherwise} \end{cases} \\
&= F_X(y) \text{ by definition of } F_X(y) \text{ being a CDF}
\end{aligned}$$

Thus $F_Y(y)=F_X(y)\ \forall\ y\in[x_0,\infty]$.