Homework 5

Moulinath Banerjee

November 11, 2020

1 The distribution of the number of sights of a jungle animal over a long period of time is given by the following logarithmic series distribution:

$$P(X=x) = -\frac{(1-p)^x}{x \log p}\ ,\ x = 1, 2, 3, \ldots\ .$$

(a) Verify that this is a proper probability distribution.
(b) What is the most likely number of sights of the jungle animal?
(c) Calculate the mean and variance of the number of sights.

2 For a random vector $\underline{X} := (X_1, X_2, \ldots, X_p)$, the m.g.f is defined as $M_{\underline{X}}(\mathbf{t}) = E(\exp(\mathbf{t}^T \underline{X}))$. As in the one-dimensional case, if $M_{\underline{X}}(\mathbf{t})$ exists for all $\|\mathbf{t}\| < \epsilon$ (for some $\epsilon > 0$), it uniquely determines the distribution of $\underline{X}$.
(a) If $X_1, X_2, \ldots, X_p$ are i.i.d. random variables, each with m.g.f $M_X(t)$, express the m.g.f. of $M_{\underline{X}}(\mathbf{t})$ in terms of $M_X(\cdot)$ where $\mathbf{t} = (t_1, t_2, \ldots, t_p)$.
(b) Find the m.g.f. of $\underline{X}$ where the X_i's are i.i.d. $N(0,1)$.
(c) Let $\underline{Y} = \mu + B\underline{X}$ where $B_{p\times p}$ is a non-singular matrix and μ is a fixed vector in $\mathbb{R}^p$. Define $\Sigma = BB^T$. Then Σ is a positive-definite matrix. Let $M_{\underline{Y}}(\mathbf{t})$ denote the m.g.f. of $\underline{Y}$.
Show that the m.g.f of $\underline{Y}$ exists for all $\mathbf{t}$ and depends only on (μ, Σ). Note that $\mu = E(\underline{Y})$ and $\Sigma = \mathrm{Cov}(\underline{Y})$, where Cov denotes the dispersion matrix. Denote by $M_{\mu,\Sigma}(\cdot)$ the m.g.f. of Y. It follows that any p dimensional random vector with this particular m.g.f must have the same distribution as Y, and in that case μ must be its mean and Σ its variance-covariance matrix.
(d) If $(\mu_1, \Sigma_1) \neq (\mu_2, \Sigma_2)$, show that $M_{\mu_1,\Sigma_1} \neq M_{\mu_2,\Sigma_2}$ and therefore the corresponding distributions are different.
(e) As (μ, Σ) varies over all possible pairs, we get the family of non-singular p–dimensional normal distributions. Use the change of variable theorem in p–dimensions to find the p.d.f. of $\underline{Y}$. (For $p = 2$ we get the family of bivariate normal distributions.)

3 Exercises from the textbook: 4.33, 4.34, 4.36, 4.40, 4.47.

4 [Supplemental exercise, will not count towards the grade] Let X be a random variable distributed on $[0, M]$ with a continuous p.d.f. $f(x)$ and distribution function $F(x)$. Assume that $f(x)$ is non-increasing on $[0, M]$, i.e. $f(u) \geq f(v)$ for $u < v$. (For purposes of visualization, you might want to draw a schematic diagram of a non-negative continuous non-increasing function on a finite interval starting at the point 0.) For this problem you will crucially need to use the fact that f is a non-increasing function.

(a) Show that $F(x) \geq x\, f(x)$ and deduce that $\mu \leq \int_0^M F(x)\, dx$ where μ is the mean of X.

(b) Show that m, the median of X (i.e. $F(m) = 1/2$) and μ are both at most $M/2$. [**Hints:** Which is greater: $P(X \leq M/2)$ or $P(X > M/2)$? Also recall that $\int_0^M (1 - F(x))\, dx = \mu$]

(c) Show that if $m = M/2$ or $\mu = M/2$ – in either case – X must follow the uniform distribution on $[0, M]$. [**Note:** In either case, you should be able to deduce that the non-increasing density f actually has to be constant.]