Tutorial n°3 – Exponential families

stoehr@ceremade.dauphine.fr

Exercise 1. Find a statistical model for the following situations

1. A production line must guarantee minimum quality requirements. The proportion θ of defective pieces produced must remain below a threshold set by the customer. A sample of pieces is analysed to determine if the quality requirements are satisfied.
2. In order to set its advertising rates, a marketing brand wants to know the time spent by a customer on a website. The brand leads a study by collecting the duration of connection of a sample.
3. A factory produces pitot tubes. The diameter of the tubes slightly differ from each other due to the production process or measurement errors. The post-production line wants to check the diameter of the tube and/or study the variation of the product from a sample.

Exercise 2. Which of the following distributions belong to the exponential family? Where appropriate, tell if the distribution is in canonical form and give its natural parameter space.

1. The normal distribution with known mean μ:
$$f(y \mid \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(y-\mu)^2}{2\sigma^2}\right].$$
2. The binomial distribution with n known:
$$f(y \mid p) = \binom{n}{y} p^y (1-p)^{n-y}.$$
3. The uniform distribution
$$f(y \mid \theta) = \theta^{-1} \mathbb{1}_{\{y \in [0,\theta]\}}.$$
4. The exponential distribution
$$f(y \mid \lambda) = \lambda e^{-\lambda y} \mathbb{1}_{\{y \geq 0\}}.$$
5. The Gamma distribution
$$f(y \mid \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-y\beta} \mathbb{1}_{\{y \geq 0\}}.$$
   a. with known shape parameter $\alpha > 0$;
   b. with known scale parameter $\beta > 0$;
6. The Gumbel distribution with known scale parameter $\phi > 0$:
$$f(y \mid \theta) = \frac{1}{\phi} \exp\left(\frac{y-\theta}{\phi} - \exp\left\{\frac{y-\theta}{\phi}\right\}\right).$$

Exercise 3. For each of the following distributions, justify if it corresponds to to a minimal exponential family and give the natural parameter space.

1. The normal distribution with unknown mean and variance.
2. The Gamma distribution with unknown shape and scale.

3. The Beta distribution:

$$f(y \mid \alpha, \beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} y^{\alpha-1}(1-y)^{\beta-1}\mathbb{1}_{\{0\leq y\leq 1\}}.$$

Exercise 4. For the following distributions, give the moment generating function of $T(X)$.

1. The Poisson distribution:

$$f(y \mid \lambda) = \frac{e^{-\lambda}\lambda^y}{y!}.$$

2. The Weibull distribution with known shape $k = 2$:

$$f(y \mid \lambda) = \frac{k}{\lambda}\left(\frac{y}{\lambda}\right)^{k-1} \exp\left[-\left(\frac{y}{\lambda}\right)^k\right]\mathbb{1}_{\{y\geq 0\}}.$$

Exercise 5. Let consider the Pareto model with known parameter $\lambda > 0$ whose density is given by

$$f(x \mid \alpha) = (\alpha-1)\lambda^{\alpha-1}x^{-\alpha}\mathbb{1}_{\{x\in[\lambda,+\infty]\}}.$$

1. Show that the Pareto model belongs to an exponential family. Is the distribution in the canonical form? Is this exponential family regular? minimal?

2. For its canonical and minimal representation, show that the natural statistic is also distributed according to an exponential family.

3. Give the first two moments of the natural statistic.

Exercise 6. Let X be a random variable whose density belongs to an exponential family with inverse normalising constant $c(\theta)$ and canonical statistic $T(X) = (T_1(X),\dots,T_d(X))$.

1. Show that for $i, j \in \{1,\dots,d\}$

$$\begin{cases} \mathbb{E}_\theta[T_i(X)] = -\frac{\partial}{\partial\theta_i}\log c(\theta), \\ \mathrm{Cov}_\theta\left[T_i(X), T_j(X)\right] = -\frac{\partial}{\partial\theta_i\partial\theta_j}\log c(\theta). \end{cases}$$

2. Use the previous result to deduce the mean and the variance of the binomial distribution with known number of trials n.