Statistical Inference and Learning- Ex-1

Yiqing Wang and Boaz Nadler

due on Wed. Nov. 25 2015

Q1 The mean height in Israel is 169cm, with standard deviation of 9cm. What is the probability of an Israeli being taller than 190cm according to Markov? Chebychev? Assuming a Gaussian distribution? Can you use Hoeffding ?

Q2

Let $X_1, \dots, X_n$ be n i.i.d. random variables from a distribution F with mean μ and finite variance σ^2. Recall that the standard error of the sample mean, namely $\sqrt{\mathbb{E}(\bar{X}_n - \mu)^2}$ decays like $1/\sqrt{n}$. Interestingly, there are cases where one can estimate population parameters (such as the mean) at a much *faster* rate. If they exist, these are known as *superefficient* estimators.

In this question we consider one such case: Let X_i be i.i.d. from a uniform distribution $U[0, \theta]$, where θ is an unknown parameter. Clearly the mean of X_i is $\mu = \theta/2$.

i) One method to estimate μ is by the sample mean $\frac{1}{n}\sum_i X_i$. What is the standard deviation of this estimator ?

ii) A different method is based on the maximum observed value of the X_i's. Let $Y_n = \max X_i$. What is $\mathbb{E}[Y_n]$ and what is $Var[Y_n]$. How can these be used to construct a different estimator for μ, and what is its error as a function of sample size n.

Q3 The Kullback - Leibler divergence between two distributions with densities p, q is defined as

$$D_{KL}[p||q] = \int p(x) \log\left(\frac{p(x)}{q(x)}\right) dx$$

1. Show that $D_{KL}[p||q] \geq 0$ and $D_{KL}[p||q] = 0$ if and only if $p = q$ almost everywhere. Hint: use the fact that
$$\log\left(\frac{p(x)}{q(x)}\right) \leq \frac{p(x)}{q(x)} - 1$$
and equality holds i.f.f. $p(x) = q(x)$.
2. Show that $D_{KL}[p||q]$ is not necessarily symmetric, that is $D_{KL}[p||q] \neq D_{KL}[q||p]$.
3. Compute $D_{KL}[p_1||p_2]$ for the following Gaussian distributions:
$$p(x) = \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} \quad (1)$$
$$q(x) = \frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-\frac{(x-\mu_2)^2}{2\sigma_2^2}} \quad (2)$$
4. Suppose p and q both have the same support (e.g. are absolutely continuous w.r.t. each other). Does this imply that their KL divergence is finite, $D_{KL}[p||q] < \infty$?

Q4 Consider a real-valued random variable X for which there is some constant $c > 0$ such that

$$\forall \lambda \in (-c, c), \qquad \mathbb{E}e^{\lambda X} < +\infty.$$

1. Show that the expectation of the absolute value of X, $\mathbb{E}|X|$, is finite.
2. Show that the function

$$f : \lambda \in (-c, c) \mapsto \mathbb{E}[e^{\lambda X}]$$

   is infinitely differentiable and its k-th derivative is

$$\forall \lambda \in (-c, c), \quad f^{(k)}(\lambda) = \mathbb{E}[X^k e^{\lambda X}].$$

3. Show that there is some constant $s > 0$ such that

$$\forall \lambda \in [-c/2, c/2], \ \ \mathbb{E}e^{\lambda(X - \mathbb{E}X)} \le e^{s\lambda^2}.$$

4. Show that as a result, we have

$$\mathbb{P}(X \ge \mathbb{E}X + t) \le \begin{cases} e^{-\frac{t^2}{4s}}, & t \in [0, sc] \\ e^{-\frac{ct}{4}}, & t \in (sc, +\infty). \end{cases}$$

5. Let $Z = \sum_{i=1}^{m} X_i^2$ where $(X_i)_{1 \le i \le m}$ are m i.i.d. Gaussian r.v. $\mathcal{N}(0, 1)$. The random variable Z follows a chi-squared distribution with m degrees of freedom, and is denoted as $Z \sim \chi^2_m$. Compute the mean and standard deviation of Z. Then, show

$$\forall t > 0, \ \ \mathbb{P}(\chi^2_m \ge m + 2mt + 2m\sqrt{t}) \le e^{-mt}.$$