14.381: Statistics

Problem Set 1

1. Let X and Y be random variables with finite variances. Show that

$$\min_{g(\cdot)} E\left(Y - g(X)\right)^2 = E\left(Y - E\left(Y \mid X\right)\right)^2,$$

where $g(\cdot)$ ranges over all functions.

2. Let X have Poisson distribution, with parameter θ

$$X \sim \text{Poisson}(\theta) \Longleftrightarrow P\left\{X = j\right\} = \frac{e^{-\theta}\theta^j}{j!} \qquad j = 0, 1, ...$$

Let $Y \sim$Poisson(λ) be independent on X.

   (a) Show that $X + Y \sim$Poisson$(\lambda + \theta)$.

   (b) Show that $X|X + Y$ is binomial with success probability $\frac{\theta}{\theta+\lambda}$.

   Note: Variable ξ has binomial distribution with success probability p and parameter n if

$$P\left\{\xi = j\right\} = \frac{n!}{j!(n-j)!}p^j(1-p)^{n-j}; \qquad j = 0, 1, \ldots, n$$

3. Show that if a sequence of random variables ξ_i converges in distribution to a constant c, then $\xi_i \xrightarrow{p} c$.

4. Let $\{X_i\}$ be independent Bernoulli (p). Then $EX_i = p$, $Var(X_i) = p(1-p)$. Let $Y_n = \frac{1}{n}\sum_{i=1}^n X_i$.

   (a) Show that $\sqrt{n}(Y_n - p) \Rightarrow N(0, p(1-p))$.

   (b) Show that for $p \neq \frac{1}{2}$ the estimated variance $Y_n(1 - Y_n)$ has the following limit behavior

$$\sqrt{n}(Y_n(1 - Y_n) - p(1-p)) \Rightarrow N(0, (1-2p)^2 p(1-p)).$$

(c) Show that for $p = \frac{1}{2}$

$$n\left[Y_n(1-Y_n) - \frac{1}{4}\right] \Rightarrow -\frac{1}{4}\chi_1^2$$

Note: χ_1^2 is a chi-square distribution with 1 degree of freedom. Let $\xi_1, \ldots, \xi_p$ be i.i.d. $N(0,1)$, then $\chi_p^2 = \sum_{i=1}^p \xi_i^2$.

Curious fact: Note that $Y_n(1-Y_n) \leq \frac{1}{4}$, that is, we always underestimate the variance for $p = \frac{1}{2}$.

(d) Prove that if (i) $\frac{\sqrt{n}}{\sigma}(\xi_n - \mu) \Rightarrow N(0,1)$ (ii) g is twice continuously differentiable: $g'(\mu) = 0$, $g''(\mu) \neq 0$, then

$$n(g(\xi_n) - g(\mu)) \Rightarrow \sigma^2 \frac{g''(\mu)}{2}\chi_1^2.$$

Note. You may assume that g has more derivatives, if it simplifies your life.

5. Let $X_1, X_2, ..., X_n \sim i.i.d.\ N(\mu, \sigma^2)$. Let us define

$$\overline{X} = \frac{1}{n}\sum_{i=1}^n X_i, s_X^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X})^2;$$

and

$$Y_j = \frac{X_j - \mu}{\sigma}; \overline{Y}_k = \frac{1}{k}\sum_{i=1}^k Y_i; s_k^2 = \frac{1}{k-1}\sum_{i=1}^k (Y_i - \overline{Y}_k)^2.$$

(a) Show that $\frac{(n-1)s_X^2}{\sigma^2} = (n-1)s_n^2$.

(b) Check that

$$(k-1)s_k^2 = (k-2)s_{k-1}^2 + \frac{k-1}{k}\left(Y_k - \overline{Y}_{k-1}\right)^2.$$

(c) Prove that if $(k-2)s_{k-1}^2 \sim \chi_{k-2}^2$ then $(k-1)s_k^2 \sim \chi_{k-1}^2$.

(d) Check that $s_2^2 \sim \chi_1^2$.

(e) Conclude that $\frac{n-1}{\sigma^2}s_X^2 \sim \chi_{n-1}^2$.

Hint: You can use the fact proved on the lecture that random variables s_k^2 and $\overline{Y}_k$ are independent.

14.381 Statistical Method in Economics
Fall 2013