THE PENNSYLVANIA STATE UNIVERSITY
Department of Economics

Economics 501 — Gallant
Final Exam Answers — Fall 2014

1. (10%) Show that the random variables $\mathcal{E}(Y|\mathcal{F}_0)$ and $[Y - \mathcal{E}(Y|\mathcal{F}_0)]$ are orthogonal in the sense that $\mathcal{E}\left\{\mathcal{E}(Y|\mathcal{F}_0)\,[Y - \mathcal{E}(Y|\mathcal{F}_0)]\right\} = 0$.

   Answer:

$$
\begin{aligned}
\mathcal{E}\left\{\mathcal{E}(Y|\mathcal{F}_0)\,[Y - \mathcal{E}(Y|\mathcal{F}_0)]\right\} &= \mathcal{E}\left\{\,\mathcal{E}\left\{\mathcal{E}(Y|\mathcal{F}_0)\,[Y - \mathcal{E}(Y|\mathcal{F}_0)]\,|\,\mathcal{F}_0\right\}\right\} \text{ iterated expectations} \\
&= \mathcal{E}\Big\{\,\mathcal{E}\Big\{\,\mathcal{E}(Y|\mathcal{F}_0)Y - \mathcal{E}(Y|\mathcal{F}_0)^2\,|\,\mathcal{F}_0\Big\}\Big\} \text{ factorization} \\
&= \mathcal{E}\Big\{\,\mathcal{E}\Big\{\,[\mathcal{E}(Y|\mathcal{F}_0)^2 - \mathcal{E}(Y|\mathcal{F}_0)^2]\,|\,\mathcal{F}_0\Big\}\Big\} \text{ linearity \& factorization} \\
&= \mathcal{E}(0)
\end{aligned}
$$

2. (15%) Let X and X_n, $n = 1, \ldots, \infty$ be random variables defined on a probability space $(\Omega, \mathcal{F}, P)$ each with range in an open set $\mathcal{X}$. Let $g(x)$ be continous on $\mathcal{X}$.

   (a) Suppose $\lim_{n\to\infty} X_n = X$ almost surely. Prove that $\lim_{n\to\infty} g(X_n) = g(X)$ almost surely.

   (b) Suppose $\lim_{n\to\infty} X_n = X$ in probability Prove that $\lim_{n\to\infty} g(X_n) = g(X)$ in probability.

   Answer:

   From lecture:

   - $g(x)$ continuous at $x \in \mathcal{X} \Leftrightarrow \lim g(x_n) = g(\lim x_n)\ \forall\ x_n$ that converges to x.
   - $X_n \xrightarrow{P} X \Rightarrow\ \exists\, X_{n_i} \ni X_{n_i} \xrightarrow{a.s.} X$
   - $X_n \xrightarrow{P} X \Leftrightarrow \forall X_{n_i}\, \exists\, X_{n_{i_j}} \ni X_{n_{i_j}} \xrightarrow{a.s.} X$

   (a) $X_n(\omega) \xrightarrow{a.s.} X(\omega)$

   $\Rightarrow \lim X_n(\omega) = \lim X(\omega)$ for $\omega \notin E$ where $P(E) = 0$

   $\Rightarrow \lim g[X_n(\omega)] = g[\lim X_n(\omega)]$ for $\omega \notin E$ where $P(E) = 0$ by second bullet above

   $\Rightarrow g(X_n) \xrightarrow{a.s.} g(X)$

(b) Let $g(X_{n_i})$ be any subsequence of $g(X_n)$.

$X_n \xrightarrow{P} X$

$\Rightarrow X_{n_i} \xrightarrow{P} X$

$\Rightarrow \exists X_{n_{i_j}} \ni X_{n_{i_j}} \xrightarrow{a.s.} X$

$\Rightarrow g(X_{n_{i_j}}) \xrightarrow{a.s.} g(X)$ by (a) above

$\Rightarrow g(X_n) \xrightarrow{P} g(X)$ by third bullet above

3. (10%) Let E_i, $i = 1, \ldots, \infty$ be events from the probability space $(\Omega, \mathcal{F}, P)$. Prove that $\sum_{i=1}^{\infty} P(E_i) < \infty$ implies $P[E_i \ i.o.] = 0$.

   Answer:

   $[E_i \ i.o.] = \bigcap_{I=1}^{\infty} \bigcup_{i=I}^{\infty} E_i \subset \bigcup_{i=I}^{\infty} E_i \ \ \forall i \Rightarrow P\left[E_i \ i.o.\right] \leq \lim_{I \to \infty} \sum_{i=I}^{\infty} P(E_i) = 0$

4. (15%) Hoeffding's inequality states that independent random variables Y_i with zero mean and bounded range, i.e., $-B < Y_i < B$, satisfy $P(|Y_1 + Y_2 + \ldots + Y_n| \geq \eta) \leq 2\exp[(-2\eta^2)/(4nB^2)]$. Define $\bar{Y_n} = (1/n) \sum_{i=1}^{n} Y_i$. Use Hoeffding's inequality and the result of Problem 3 to prove that $P[|\bar{Y_n}| > n^{-\frac{1}{2}+\tau} \ i.o.] = 0$ for every $\tau > 0$.

   Answer:

   $P(|\bar{Y_n}| \geq n^{-\frac{1}{2}+\tau}) = P(|Y_1 + \ldots + Y_n| \geq n^{\frac{1}{2}+\tau}) \leq 2\exp(-\frac{1}{2}n^{2\tau}/B^2)$

   $\sum \exp(-\frac{1}{2}n^{2\tau}/B^2) < \infty \Rightarrow P[|\bar{Y_n}| \geq n^{-\frac{1}{2}+\tau} \ i.o.] = 0$

5. (15%) Let X and Z be measurable functions mapping $(\Omega, \mathcal{F})$ to $(\mathbb{R}^d, \mathcal{B})$, where $\mathcal{B}$ is the collection of Borel subsets of $\mathbb{R}^d$. Let $\mathcal{X}$ be the range of X, that is, $\mathcal{X} = X(\Omega)$. Let $\mathcal{F}_0 = X^{-1}(\mathcal{B})$, that is, $\mathcal{F}_0 = \{F : F = X^{-1}(B), B \in \mathcal{B}\}$. Assume Z is measurable $(\Omega, \mathcal{F}_0)$. Prove that there exists a function g mapping $\mathcal{X}$ into $\mathbb{R}^d$ such that $Z(\omega) = g[X(\omega)]$.

   Answer:

   Let $F_z = Z^{-1}(\{z\})$. $F_z \in \mathcal{F}_0 \Rightarrow \exists \ B_z \in \mathcal{B} \ni F_z = X^{-1}(B_z)$. Let $g(x) = \sup_{z \in \mathcal{Z}} z I_{B_z \cap \mathcal{X}}(x)$. Then

$$g[X(\omega)] = \sup_{z \in \mathcal{Z}} z I_{B_z \cap \mathcal{X}}[X(\omega)] = \sup_{z \in \mathcal{Z}} z I_{X^{-1}[B_z]}(\omega) = \sup_{z \in \mathcal{Z}} z I_{F_z}(\omega) = Z(\omega)$$

   The last equality is because the F_z are disjoint, Z is constant on F_z, and that constant is z.

6. (10%)

   (a) State Chebishev's inequality.

   (b) Prove Chebishev's inequality for a general random variable, that is, do not assume that the random variable has a density.

   Answer:

   (a) $P(|X-\mu| > \epsilon) \leq \frac{Var(X)}{\epsilon^2}$

   (b) Let $Y = X - \mu$. Then $\epsilon^2 P(|X-\mu| > \epsilon) = \epsilon^2 P(|Y| > \epsilon) = \int_{-\infty}^{-\epsilon} \epsilon^2 dP_Y + \int_{\epsilon}^{\infty} \epsilon^2 dP_Y \leq \int_{-\infty}^{-\epsilon} Y^2 dP_Y + \int_{-\epsilon}^{\epsilon} Y^2 dP_Y + \int_{\epsilon}^{\infty} Y^2 dP_Y = \mathrm{Var}(Y) = \mathrm{Var}(X)$

7. (10%) The Metropolis-Hastings algorithm is as follows:

   - Proposal density: $T(\theta_{here}, \theta_{there})$
   - Posterior: $f(\theta|x_1,\ldots,x_n) = [\prod_{i=1}^n f(x_i|\theta)]\, \pi(\theta) \,/ \int [\prod_{i=1}^n f(x_i|\theta)]\, \pi(\theta)\, d\theta$
   - Propose: Draw θ_{prop} from $T(\theta_{old}, \theta)$
   - Compute: $\alpha = ?$
   - Put θ_{new} to θ_{prop} with probability α (accept)
   - Put θ_{new} to θ_{old} with probability $1-\alpha$ (reject)

   (a) Write the formula for α.

   (b) Why is it not necessary to know the normalization factor $\int [\prod_{i=1}^n f(x_i|\theta)] \pi(\theta)\, d\theta$ in order use the Metropolis-Hastings algorithm.

   (c) What determines the rejection rate of the chain?

   (d) What is the theoretically best choice of proposal density.

   Answer:

   (a) $\alpha = \min\left[1, \frac{f(\theta_{prop}|x_1,\ldots,x_n) T(\theta_{prop}, \theta_{old})}{f(\theta_{old}|x_1,\ldots,x_n) T(\theta_{old}, \theta_{prop})}\right]$

   (b) The normalization factor does not involve θ and therefore cancels out in the computation of α.

   (c) The scale of the proposal density.

   (d) Iid draws from the posterior.

8. (15%) Let $x_1, x_2, \ldots, x_n$ be independent $n(x_i \,|\, \mu, \sigma^2)$ random variables. Define

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

$$U = \begin{pmatrix} \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \cdots & \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 & \cdots & 0 & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} & 0 & \cdots & 0 & 0 \\ & & & & \vdots & & \\ \frac{1}{\sqrt{n(n-1)}} & \frac{1}{\sqrt{n(n-1)}} & \frac{1}{\sqrt{n(n-1)}} & \frac{1}{\sqrt{n(n-1)}} & \cdots & \frac{1}{\sqrt{n(n-1)}} & -\frac{n-1}{\sqrt{n(n-1)}} \end{pmatrix}$$

and $z = Ux$.

You may use without proof the fact that $UU' = U'U = I$.

(a) Show that $z_1 = \sqrt{n}\bar{x}$.

(b) Show that $\sum_{i=2}^{n} z_i^2 = (n-1)s^2$.

(c) Show that the density of z is

$$f(z) = n(z_1 \,|\, \sqrt{n}\mu,\, \sigma^2) \times \prod_{i=2}^{n} n(z_i \,|\, 0,\, \sigma^2)$$

(d) Why does 8c imply that $\bar{x}$ and s^2 are independent?

Answer:

(a) $z_1 = \frac{1}{\sqrt{n}}\mathbf{1}'x = \sqrt{n}\frac{1}{n}\sum_{i=1}^{n} x_i$

(b) $(n-1)s^2 = x'x - n\bar{x}^2 = x'U'Ux - n\bar{x}^2 = \sum_{i=1}^{n} z_i^2 - z_1^2$

(c) A linear transform of the $N(\mu, \sigma^2)$ is the multivariate normal with mean $\mathcal{E}z = (\sqrt{n}\mu, 0, \ldots, 0)'$ and variance $\mathcal{C}(z, z') = U(\sigma^2 I)U' = \sigma^2 I$

(d) Density factors.