Econ 240A Econometrics **Fall 2018**

Problem Set 2 Solutions

Due date: Sept. 17, 2018 *Fengshi Niu*

1. Probability integral transform

Since $F:\mathbb{R}\to[0,1]$ is continuous and strictly increasing, its inverse, $F^{-1}:(0,1)\to\mathbb{R}$, exists and is also strictly increasing. This implies $x\le y \Leftrightarrow F^{-1}(x)\le F^{-1}(y)$. For any $y\in(0,1)$,

$$\begin{aligned} F_Y(y) &= P(Y\le y) = P\left(F\left(X\right)\le y\right) = P\left(F^{-1}\left(F\left(X\right)\right)\le F^{-1}\left(y\right)\right) \\ &= P\left(X\le F^{-1}\left(y\right)\right) \\ &= F\left(F^{-1}\left(y\right)\right) \\ &= y. \end{aligned}$$

This shows

$$F_Y(y)=\begin{cases} 0 & x\le 0 \\ y & 0<y<1 \\ 1 & y\ge 1 \end{cases}.$$

Hence, $Y\sim \mathrm{U}[0,1]$.

2. Inverse transform sampling

Since $F:\mathbb{R}\to[0,1]$ is continuous and strictly increasing, its inverse, $F^{-1}:(0,1)\to\mathbb{R}$, is well defined. Also notice $x\le y\Leftrightarrow F(x)\le F(y)$. For any $x\in\mathbb{R}$,

$$\begin{aligned} F_X(x) &= P(X\le x) = P\left(F^{-1}\left(Y\right)\le x\right) = P\left(F\left(F^{-1}\left(Y\right)\right)\le F\left(x\right)\right) \\ &= P\left(Y\le F\left(x\right)\right) \\ &= F(x). \end{aligned}$$

Hence, $X\sim F$.

3. Moments and moment generating functions

(a) **Method 1**

For any $0<l<r$, $|x|^l\le|x|^r$, if $|x|\ge1$; $|x|^l<1$, if $|x|<1$. This suggests for any $x\in\mathbb{R}$,

$$|x|^l\le \mathbf{1}(|x|<1)\cdot 1+\mathbf{1}(|x|\ge 1)|x|^r<1+|x|^r.$$

Hence, $|X|^l<1+|X|^r$. Take expectation on both sides.

$$\mathbb{E}(|X|^l)<1+\mathbb{E}(|X|^r).$$

If the RHS is finite, the LHS must be finite.

Method 2
For any $0 < l < r$, the function $g : [0, \infty) \to \mathbb{R}$, $g(x) = x^{\frac{l}{r}}$, is a concave function on $[0, \infty)$. Use Jensen's inequality for r.v. $Y := |X|^r$ and function g.

$$\mathbb{E}(|X|^l) = \mathbb{E}\left((|X|^r)^{\frac{l}{r}}\right) = \mathbb{E}\left(g(Y)\right) \leq g\left(\mathbb{E}\left(Y\right)\right) = \left(\mathbb{E}\left(|X|^r\right)\right)^{\frac{l}{r}}.$$

If the RHS is finite, the LHS must be finite.

(b) $X \sim \mathrm{N}(0, 1)$.

$$\begin{aligned}
M_X(t) &= \mathbb{E}(e^{tX}) \\
&= \int e^{tx} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\
&= \int \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-t)^2}{2}} dx \cdot e^{\frac{t^2}{2}} \\
&= e^{\frac{t^2}{2}} \\
\mathbb{E}(X) &= \frac{d}{dx} M_X(t)|_{t=0} = te^{\frac{t^2}{2}}|_{t=0} = 0 \\
\mathbb{E}(X^2) &= \frac{d^2}{dx^2} M_X(t)|_{t=0} = (t^2+1)e^{\frac{t^2}{2}}|_{t=0} = 1 \\
\mathbb{E}(X^3) &= \frac{d^3}{dx^3} M_X(t)|_{t=0} = (t^3+3t)e^{\frac{t^2}{2}}|_{t=0} = 0 \\
\mathbb{E}(X^4) &= \frac{d^4}{dx^4} M_X(t)|_{t=0} = (t^4+6t^2+3)e^{\frac{t^2}{2}}|_{t=0} = 3.
\end{aligned}$$

4. Covariance

(a) $(X, Y)' \sim U\left((-1, 1) \times (-1, 1)\right)$ means

$$f_{X,Y}(x, y) = \begin{cases} \frac{1}{4} & (x, y)' \in (-1, 1) \times (-1, 1) \\ 0 & \text{otherwise} \end{cases}.$$

$$\begin{aligned}
\mathrm{Cov}(X, Y) &= \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) \\
&= \int\int xy \cdot f_{X,Y}(x, y) dx dy - \int\int x \cdot f_{X,Y}(x, y) dx dy \cdot \int\int y \cdot f_{X,Y}(x, y) dx dy \\
&= 0 - 0 \cdot 0 = 0 \qquad \text{by symmetry.}
\end{aligned}$$

X and Y are uncorrelated. Moreover, since

$$f_X(x) = \begin{cases} \frac{1}{2} & -1 < x < 1 \\ 0 & \text{otherwise} \end{cases} \qquad f_Y(y) = \begin{cases} \frac{1}{2} & -1 < y < 1 \\ 0 & \text{otherwise} \end{cases},$$

$f_{X,Y}(x, y) = f_X(x) \cdot f_Y(y)$ for any $(x, y)' \in \mathbb{R}^2$. This implies X and Y are independent.

(b) Denote the ball $\{(x, y)' \in \mathbb{R}^2 : x^2+y^2 \leq 1\}$ by $B\left((0, 0)', 1\right)$. $(X, Y)' \sim U\left(B\left((0, 0)', 1\right)\right)$

means

$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{\pi} & (x,y)' \in B\left((0,0)', 1\right) \\ 0 & \text{otherwise} \end{cases}.$$

$$\begin{aligned}
\operatorname{Cov}(X,Y) &= \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) \\
&= \int\int xy \cdot f_{X,Y}(x,y)dxdy - \int\int x \cdot f_{X,Y}(x,y)dxdy \cdot \int\int y \cdot f_{X,Y}(x,y)dxdy \\
&= 0 - 0 \cdot 0 = 0 \qquad \text{by symmetry.}
\end{aligned}$$

X and Y are uncorrelated. However, since

$$f_X(x) = \begin{cases} \frac{2\sqrt{1-x^2}}{\pi} & -1 < x < 1 \\ 0 & \text{otherwise} \end{cases} \qquad f_Y(y) = \begin{cases} \frac{2\sqrt{1-y^2}}{\pi} & -1 < y < 1 \\ 0 & \text{otherwise} \end{cases},$$

$f_{X,Y}(x,y) \neq f_X(x) \cdot f_Y(y)$ for a set in $\mathbb{R}^2$ with positive measure. This implies X and Y are not independent.

Remark. *Other arguments also work. For example, the conditional distribution of $Y|X = x$ clearly depends on x. This implies X and Y are not independent. Or the cdf of the joint isn't equal to the product of the marginals'.*

(c) Since $g(\cdot)$ and $h(\cdot)$ are non-decreasing functions on $\mathbb{R}$,

$$(g(x) - g(y))\,(h(x) - h(y)) \geq 0, \ \forall x, y \in \mathbb{R}.$$

Consider random variables $X_1, X_2 \overset{\text{iid}}{\sim} F$. We have

$$\begin{aligned}
&(g(X_1) - g(X_2))\,(h(X_1) - h(X_2)) \geq 0 \\
\Longrightarrow &\mathbb{E}\left[(g(X_1) - g(X_2))\,(h(X_1) - h(X_2))\right] \geq 0.
\end{aligned}$$

Notice

$$\begin{aligned}
0 &\leq \mathbb{E}\left[(g(X_1) - g(X_2))\,(h(X_1) - h(X_2))\right] \\
&= \mathbb{E}\left[g(X_1)h(X_1)\right] - \mathbb{E}\left[g(X_1)h(X_2)\right] - \mathbb{E}\left[g(X_2)h(X_1)\right] + \mathbb{E}\left[g(X_2)h(X_2)\right] \\
&= \mathbb{E}\left[g(X)h(X)\right] - \mathbb{E}[g(X)]\mathbb{E}[h(X)] - \mathbb{E}[g(X)]\mathbb{E}[h(X)] + \mathbb{E}\left[g(X)h(X)\right] \\
&= 2\Big(\mathbb{E}\left[g(X)h(X)\right] - \mathbb{E}[g(X)]\mathbb{E}[h(X)]\Big) \\
&= 2\operatorname{Cov}\left(g(X), h(X)\right) \\
\Longrightarrow &\operatorname{Cov}\left(g(X), h(X)\right) \geq 0.
\end{aligned}$$

5. The gamma distribution

(a)

$$\begin{aligned}
\Gamma(x+1) &= \int_0^\infty t^x e^{-t} dt \\
&= -\int_0^\infty t^x de^{-t} \\
&= -(t^x e^{-t}|_0^\infty - \int_0^\infty e^{-t} dt^x) \\
&= x\int_0^\infty t^{x-1} e^{-t} dt \\
&= x\Gamma(x).
\end{aligned}$$

Since $\Gamma(1) = \int_0^\infty e^{-t}dt = -e^{-t}|_0^\infty = 1 = 0!$,

$$\begin{aligned}
\Gamma(n+1) &= n\Gamma(n) \\
&= n\cdot(n-1)\Gamma(n-2) \\
&\dots \\
&= n\cdot(n-1)\cdots 2\cdot 1\Gamma(1) \\
&= n! && \text{for any } n = 1, 2, \ldots.
\end{aligned}$$

(b) *

(c) First, $p(x) \geq 0$ for any $x \in \mathbb{R}$. Second,

$$\begin{aligned}
\int p(x)dx &= \int_0^\infty \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta} dx \\
&= \int_0^\infty \frac{1}{\Gamma(\alpha)} t^{\alpha-1} e^{-t} dt && \text{by change of variable } t = x/\beta \\
&= \frac{\Gamma(\alpha)}{\Gamma(\alpha)} \\
&= 1.
\end{aligned}$$

This shows $p(\cdot)$ is a pdf.

(d) *

(e) *

(f) *

6. Best linear predictor

To get the minimizer (α^*, β^*) of the criteria function $h(\alpha,\beta) := \mathbb{E}\left[(Y - (\beta X + \alpha))^2\right]$, we look at the derivatives of h.

$$\begin{aligned}
\frac{\partial}{\partial\alpha} h(\alpha,\beta) &= -2[\mathbb{E}(Y) - \beta\mathbb{E}(X) - \alpha] \\
\frac{\partial}{\partial\beta} h(\alpha,\beta) &= -2\mathbb{E}[(Y - \beta X - \alpha)X]
\end{aligned}$$

The first-order condition says the two partial derivatives are 0 at the minimizer.

$$
\begin{aligned}
&-2[\mathbb{E}(Y) - \beta^*\mathbb{E}(X) - \alpha^*] = 0 \\
&-2\mathbb{E}[(Y - \beta^* X - \alpha^*)X] = 0.
\end{aligned}
$$

Solve this system of linear equations. The first equation implies $\alpha^* = \mathbb{E}\,(Y) - \beta^*\mathbb{E}\,(X)$. Plug it into the second equation. When $\mathbb{E}(X^2) - (\mathbb{E}X)^2 = \mathrm{Var}(X) > 0$ (the following discussion assume $\mathrm{Var}(X) > 0$),

$$
\begin{aligned}
\beta^* &= \frac{\mathbb{E}[(Y - \mathbb{E}Y)X]}{\mathbb{E}[(X - \mathbb{E}X)X]} \\
&= \frac{\mathrm{Cov}\,(X, Y)}{\mathrm{Var}\,(X)}.
\end{aligned}
$$

Next, let's check the second-order condition.

$$
H(\alpha, \beta) = \begin{pmatrix} \frac{\partial^2}{\partial\alpha^2} h(\alpha, \beta) & \frac{\partial^2}{\partial\beta\partial\alpha} h(\alpha, \beta) \\ \frac{\partial^2}{\partial\alpha\partial\beta} h(\alpha, \beta) & \frac{\partial^2}{\partial\beta^2} h(\alpha, \beta) \end{pmatrix} = 2 \begin{pmatrix} 1 & \mathbb{E}(X) \\ \mathbb{E}(X) & \mathbb{E}(X^2) \end{pmatrix}.
$$

The Hessian doesn't depend on (α, β) and is positive definite; because for any $(a, b)' \in \mathbb{R}^2$, $(a, b)' \neq (0, 0)'$, the quadratic form

$$
\begin{pmatrix} a & b \end{pmatrix} \begin{pmatrix} 1 & \mathbb{E}(X) \\ \mathbb{E}(X) & \mathbb{E}(X^2) \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = a^2 + 2ab\mathbb{E}(X) + b^2\mathbb{E}(X^2) = \mathbb{E}[(a + bX)^2] > 0.
$$

This implies the function h is strictly convex on $\mathbb{R}^2$ and the solution of the first-order condition, $(\hat{\alpha}, \hat{\beta})$, is the unique global minimizer. Our derivation shows $\frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(X)}(X - \mathbb{E}(X)) + \mathbb{E}(Y)$ is the unique best linear predictor of Y using X when $\mathrm{Var}(X) > 0$. When $\mathrm{Var}(X) = 0$, X is a constant. h is weakly convex and the minimizer is not unique. Any (α, β) s.t. $\alpha + \beta\mathbb{E}(X) - \mathbb{E}(Y) = 0$ is a minimizer.

7. Kullback-Leibler divergence

Consider a random variable X with density p. Define random variable $Y := \frac{q(X)}{p(X)}$. Y is alway positive. To apply Jensen's inequality, consider the convex function $g : (0, \infty) \to \mathbb{R},\ g(x) = -\log(x)$.

$$
\begin{aligned}
-\int \log\,(\frac{q(x)}{p(x)}) p(x)dx &= \mathbb{E}(g(Y)) \\
&\geq g(\mathbb{E}(Y)) && \text{(by Jensen's inequality)} \\
&= -\log\,(\int \frac{q(x)}{p(x)} \cdot p(x)dx) \\
&= -\log\,(\int q(x)dx) \\
&= -\log(1) \\
&= 0.
\end{aligned}
$$

Since g is strictly convex, Jensen's inequality achieves equality if and only if Y is a constant a.s.. Suppose $Y = c$ a.s.. Notice $c = \mathbb{E}(Y) = \int \frac{q(x)}{p(x)} \cdot p(x)dx = 1$. Hence $Y = 1$ a.s., which means $p(X) = q(X)$ a.s.. This is just saying $p(x) = q(x)$ almost surely with respect to p.