Pre-Semester Course Statistics 2019

Tutor: Nan Hu [1]

Exercises

In case you need more exercises on the topics covered in class, consider those at the end of chapters 2 and 3 in Stock/Watson and those at the end of Appendices B and C in Wooldridge.

Exercise 1

Consider the following PMF, where k is a constant:

$$f_X(x) = \begin{cases} k - 2^{-1-x} & \text{if } x = 0, 1, 2 \\ 0 & \text{otherwise} \end{cases}$$

a) Calculate k.

b) Calculate $P(X = 1|X \geq 1)$.

Exercise 2

Assume X and Y are discrete RVs. Show that if $X \perp\!\!\!\perp Y$, one can write $F_{X,Y}(x, y) = F_X(x)F_Y(y)$.

Exercise 3

Show that $E[X + Y] = E[X] + E[Y]$ for continuous RVs X and Y.

Exercise 4

Suppose X follows a Poisson distribution with parameter λ, i.e. $X \sim Po(\lambda)$

a) Show that $E[X] = \lambda$.

b) Show that $Var[X] = \lambda$.

Exercise 5

Suppose $A = \sqrt{B}C$ with $C \sim N(0, 1)$, $B \sim arbitrary$ (but $E[B^2]$ exists and is finite) and $B \perp\!\!\!\perp C$.

a) Describe (no formulas required) why $E[C^4] = 3$.

b) Comparing with formula (27) in the slides, argue why $\kappa_A = \frac{E[A^4]}{Var[A]^2}$ in this case.

c) Show that $\kappa_A \geq 3$.

[1] Materials contributed by Sebastian Schreiber.

Exercise 6

Let X and Y be arbitrary RVs and let $Z \equiv E[X|Y]$. Furthermore assume $E[|X|] < \infty$. Show that $E[|Z|] < \infty$.

Hint: Jensen's inequality (you might have seen this in Real Analysis) reads as $f(E[X]) \leq E[f(X)]$ for convex functions and as $f(E[X]) \geq E[f(X)]$ for concave functions.

Exercise 7

Suppose $\begin{pmatrix} u_t \\ v_t \end{pmatrix} \overset{iid}{\sim} N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_u^2 & \sigma_{uv} \\ \sigma_{uv} & \sigma_v^2 \end{pmatrix}\right)$

Further let

$$x_t \equiv \sum_{j=0}^{\infty} a^j u_{t-j} + y_t$$
$$y_t \equiv \sum_{j=0}^{\infty} a^j v_{t-j}$$

where $|a| < 1$.

Show that $Cov(x_t, y_t) = \frac{\sigma_v^2 + \sigma_{uv}}{1-a^2}$.

Exercise 8

Assume that a stock price at time t, S_t, is modelled as follows:

$$ln(S_t) = ln(S_0) + \mu_S t + \sigma_S W_t$$

where the initial stock price S_0, the stock drift μ_S and the stock volatility σ_S are constants and $W_t \sim N(0, t)$.

a) Is $ln(S_t)$ are RV?

b) Is S_t are RV?

c) Is S_t iid?

d) Calculate $E[S_t]$.

e) Calculate $Var[S_t]$.

Exercise 9: Right or Wrong?

a) Any RV has a PDF.

b) For a RV, there is only one sample mean but there are multiple expected values.

c) If X is normally distributed, and $Y \equiv e^X$, Y is lognormally distributed.

d) The expectation operator is invariant to nonlinear transformations.

e) The first centralized moment is always equal to 0.

f) The CLT says that the sample mean of any well-behaved (i.e. at least the first two moments exist and are finite) distribution follows exactly a normal distribution.

Exercise 10: Patients with glaucoma in one eye

The following data (Ehlers,N., *Acta Opthalmologica*, **48**) give corneal thicknesses inmicrons for patients with one glaucomatous eye and one normal eye.

Table 4.1 Glaucoma in one eye

Corneal thickness		
Glaucoma	Normal	Difference
488	484	4
478	478	0
480	492	-12
426	444	-18
440	436	4
410	398	12
458	464	-6
460	476	-16

(1) Is there a difference in corneal thickness between the eyes? Build up your null hypothesis and conduct upper tail test, down tail test and two-sides test to answer the question by hand. Note: the one-side critical value of t distribution with degree of freedom 7 at the level of 0.05 is 1.8946, and the two-side critical value is 2.3646.

(2) Answer the question use p-values instead of critical values.

(3) Confirm you results with the result of Stata command *ttest*.

Solutions

Exercise 1

a) Since this is a PMF, we have by formula (3) in the slides

$$P(X = x) = f_X(x)$$

To find out k, we use formula (2) in the slides, namely that the probabilities must sum up to one. That is,

$$\begin{aligned} k - 2^{-1-0} + k - 2^{-1-1} + k - 2^{-1-2} &\overset{!}{=} 1 \\ k &= \frac{5}{8} \end{aligned}$$

b) By formula (32) in the slides,

$$P(X = 1|X \geq 1) = \frac{P(X = 1 \cap X \geq 1)}{P(X \geq 1)} = \frac{P(X = 1)}{P(X \geq 1)} = \frac{P(X = 1)}{P(X = 1) + P(X = 2)}$$

Using $k = \frac{5}{8}$, we find

$$\begin{aligned} P(X = 1) &= \frac{5}{8} - 2^{-1-1} = \frac{3}{8} \\ P(X = 2) &= \frac{5}{8} - 2^{-1-2} = \frac{1}{2} \end{aligned}$$

and therefore

$$P(X = 1|X \geq 1) = \frac{\frac{3}{8}}{\frac{3}{8} + \frac{1}{2}} = \frac{3}{7}$$

Exercise 2

By formula (29) in the slides,

$$\begin{aligned} F_{X,Y}(x, y) &= \sum_{i|x_i \leq x} \sum_{j|y_j \leq y} f_{X,Y}(x_i, y_i) \qquad |\text{use (36)} \\ &= \sum_{i|x_i \leq x} \sum_{j|y_j \leq y} f_X(x_i) f_Y(y_i) \qquad |f_X(x_i) \text{ can be taken out of the } y\text{-sum running over } j \\ &= \underbrace{\sum_{i|x_i \leq x} f_X(x_i)}_{=F_X(x)} \underbrace{\sum_{j|y_j \leq y} f_Y(y_i)}_{=F_Y(y)} \\ &= F_X(x) F_Y(y) \end{aligned}$$

Exercise 3

$$
\begin{aligned}
E[X+Y] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x+y)f_{X,Y}(x,y)dxdy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xf_{X,Y}(x,y)dxdy + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}yf_{X,Y}(x,y)dxdy \\
&= \int_{-\infty}^{\infty} x \underbrace{\int_{-\infty}^{\infty} f_{X,Y}(x,y)dy}_{=f_X(x)\, by\, (30)} dx + \int_{-\infty}^{\infty} y \underbrace{\int_{-\infty}^{\infty} f_{X,Y}(x,y)dx}_{=f_Y(y)\, by\, (30)} dy \\
&= \underbrace{\int_{-\infty}^{\infty} xf_X(x)dx}_{=E[X]} + \underbrace{\int_{-\infty}^{\infty} yf_Y(y)dy}_{=E[Y]} \\
&= E[X]+E[Y]
\end{aligned}
$$

Note the following: Since both intervals run from $-\infty$ to ∞, we can change the order of integration. Furthermore, we can of course e.g. take x out of the 'y-integral'.

Exercise 4

a)

$$
\begin{aligned}
E[X] &= \sum_x xf_X(x,\lambda) \\
&= \sum_{k=0}^{\infty} k\frac{e^{-\lambda}\lambda^k}{k!} && \mid \text{one can let the sum start at k=1 since for k=0 we get 0} \\
&= \sum_{k=1}^{\infty} k\frac{e^{-\lambda}\lambda^k}{k!} && \mid \text{factor out } e^{-\lambda}\lambda \\
&= e^{-\lambda}\lambda\sum_{k=1}^{\infty} k\frac{\lambda^{k-1}}{k!} && \mid \text{use } \frac{k}{k!} = \frac{k}{k\cdot(k-1)!} = \frac{1}{(k-1)!} \\
&= e^{-\lambda}\lambda\sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} && \mid \text{shift index by 1 (alternatively, let } j \equiv k-1) \\
&= e^{-\lambda}\lambda \underbrace{\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}}_{=e^{\lambda}\text{by exponential series}} = \lambda e^{-\lambda}e^{\lambda} = \lambda
\end{aligned}
$$

See `here` for the derivation of the exponential series. You definitely should know this series and the geometric series used below (but not their proofs of course).

b) The proof works to a large extent like the one in a).

Noting that $Var[X] = E[X^2] - E[X]^2 \overset{a)}{=} E[X^2] - \lambda^2$, we calculate $E[X^2]$:

$$
\begin{aligned}
E[X^2] &= \sum_{k=0}^{\infty} k^2 \frac{e^{-\lambda}\lambda^k}{k!} && \mid \text{one can let the sum start at k=1 since for k=0 we get 0}\\
&= \sum_{k=1}^{\infty} k^2 \frac{e^{-\lambda}\lambda^k}{k!} && \mid \frac{k^2}{k!} = \frac{k}{(k-1)!}\\
&= \sum_{k=1}^{\infty} k \frac{e^{-\lambda}\lambda^k}{(k-1)!} && \mid \text{factor out } e^{-\lambda}\lambda\\
&= e^{-\lambda}\lambda \sum_{k=1}^{\infty} k \frac{\lambda^{k-1}}{(k-1)!} && \mid \text{use } k = (k-1)+1\\
&= e^{-\lambda}\lambda \left(\sum_{k=1}^{\infty} (k-1) \frac{\lambda^{k-1}}{(k-1)!} + \underbrace{\sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!}}_{=e^{\lambda}} \right) && \mid \text{use } \frac{k-1}{(k-1)!} = \frac{1}{k-2}, \text{factor out } \lambda\\
&= e^{-\lambda}\lambda \left(\lambda \underbrace{\sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!}}_{=e^{\lambda}} + e^{\lambda} \right)\\
&= e^{-\lambda}\lambda(\lambda e^{\lambda} + e^{\lambda})\\
&= \lambda^2 + \lambda
\end{aligned}
$$

Therefore,

$$Var[X] = E[X^2] - \lambda^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

Exercise 5

a) Since C is (standard) normally distributed, i.e. normally distributed with parameters $\mu_C = 0$ and $\sigma_C^2 = 1$, we know the entire distribution of C and therefore all moments. The Kurtosis is the 4th centralised moment and known to be 3 for any normal distribution.
(If you want to check this, calculate $E[C^4] = \int_{-\infty}^{\infty} C^4 f_C(C)dC$, which, either using integration by parts or plugging in (71), i.e. the PDF of a normal, and solving the integral yields $E[C^4] = 3$.)
b) Since $\mu_A = E[\sqrt{B}C] = E[\sqrt{B}]\underbrace{E[C]}_{=0} + \underbrace{Cov[\sqrt{B}, C]}_{=0} = 0$, we end with $\kappa_A = \frac{E[A^4]}{Var[A]^2}$. Note that $Cov[\sqrt{B}, C] = 0$ since, as evoked on slide 55 , independence implies a covariance of 0.

c)

$$
\begin{aligned}
E[A^4] &= E[B^2C^4] \overset{B \perp\!\!\!\perp C}{=} E[B^2]E[C^4] \overset{E[C^4]=3}{=} 3E[B^2] \\
Var[A] &= E[(\sqrt{B}C)^2] - (E[\sqrt{B}C])^2 \overset{B \perp\!\!\!\perp C}{=} E[B]\underbrace{E[C^2]}_{=1} - (E[\sqrt{B}]\underbrace{E[C]}_{=0})^2 = E[B] \\
\Rightarrow \kappa_A &= \frac{E[A^4]}{Var[A]^2} = 3\frac{E[B^2]}{E[B]^2}
\end{aligned}
$$

Now since any Variance must be greater or equal to 0,

$$
Var[B] = E[B^2] - E[B]^2 \geq 0 \Leftrightarrow \frac{E[B^2]}{E[B]^2} \geq 1
$$

and therefore

$$
\kappa_A = \frac{E[A^4]}{Var[A]^2} = 3\frac{E[B^2]}{E[B]^2} \geq 3
$$

Exercise 6

First note that Jensen's Inequality (JI) in our case says that $|E[X|Y]| \leq E[|X||Y]$ since $f(x) = |x|$ and therefore $f(x) = E[|X|]$ are convex functions.

$$
E[|Z|] \overset{Z=E[X|Y]}{=} E[|E[X|Y]|] \overset{JI}{\leq} E[E[|X||Y]] \overset{LIE}{=} E[|X|] \overset{Assumption}{\leq} \infty
$$

Exercise 7

Throughout the exercise, we will use the variance/covariance rules on slide 56 as well as some geometric series.

$$
Cov[x_t, y_t] = Cov\left[\sum_{j=0}^{\infty} a^j u_{t-j} + y_t, y_t\right] = \underbrace{Cov\left[\sum_{j=0}^{\infty} a^j u_{t-j}, \sum_{j=0}^{\infty} a^j v_{t-j}\right]}_{(ii)} + \underbrace{Var[y_t]}_{(i)}
$$

(i) Let's first calculate the variance of y_t

$$
\begin{aligned}
Var[y_t] &= Var\left[\sum_{j=0}^{\infty} a^j v_{t-j}\right] \\
&= Var[v_t + av_{t-1} + a^2 v_{t-2} + \dots]
\end{aligned}
$$

By (46), the variance of a sum is the sum of the variance plus the covariances. Since v_t is iid, $Cov[v_t, v_s] = 0 \, \forall \, t \neq s$ (e.g. v_t is independent of v_{t-1} and independence implies covariance) and therefore

$$
\begin{aligned}
Var[y_t] &= Var[v_t + av_{t-1} + a^2 v_{t-2} + ..] \\
&= \underbrace{Var[v_t]}_{=\sigma_v^2} + 2a \underbrace{Cov[v_t, v_{t-1}]}_{=0} + a^2 \underbrace{Var[v_{t-1}]}_{=\sigma_v^2} + .. \\
&= \sigma_v^2 + a^2\sigma_v^2 + a^4\sigma_v^2 + \dots \\
&= \sigma_v^2 \sum_{j=0}^{\infty} (a^2)^j
\end{aligned}
$$

Since $|a| < 1$ and therefore $|a^2| < 1$, we can further simplify this result using a geometric series: $\sum_{i=0}^{\infty} k^i = \frac{1}{1-k}$ if $|k| < 1$. In our case, this yields

$$
\begin{aligned}
Var[y_t] &= \sigma_v^2 \sum_{j=0}^{\infty} (a^2)^j \\
&= \sigma_v^2 \frac{1}{1-a^2}
\end{aligned}
$$

(ii) Now for the covariance, we proceed analogously:

$$
Cov\left[\sum_{j=0}^{\infty} a^j u_{t-j}, \sum_{j=0}^{\infty} a^j v_{t-j}\right] = Cov[u_t + au_{t-1} + a^2 u_{t-2} + \dots, v_t + av_{t-1} + a^2 v_{t-2} + \dots]
$$

Again, since u_t and v_t each are iid, $Cov[u_t, v_s] = 0 \,\forall\, t \neq s$, $Cov[u_t, v_s] = \sigma_{uv} \,\forall\, t = s$ and therefore

$$
\begin{aligned}
Cov\left[\sum_{j=0}^{\infty} a^j u_{t-j}, \sum_{j=0}^{\infty} a^j v_{t-j}\right] &= Cov[u_t + au_{t-1} + a^2 u_{t-2} + \dots, v_t + av_{t-1} + a^2 v_{t-2} + \dots] \\
&= \underbrace{Cov[u_t, v_t]}_{=\sigma_{uv}} + a \underbrace{Cov[u_{t-1}, v_t]}_{=0} + a^2 \underbrace{Cov[u_{t-1}, v_{t-1}]}_{=\sigma_{uv}} + \dots \\
&= \sigma_{uv} + a^2\sigma_{uv} + a^4\sigma_{uv} + \dots \\
&= \sigma_{uv} \sum_{j=0}^{\infty} (a^2)^j \\
&= \sigma_{uv} \frac{1}{1-a^2}
\end{aligned}
$$

Therefore,

$$
Cov[x_t, y_t] = Cov\left[\sum_{j=0}^{\infty} a^j u_{t-j}, \sum_{j=0}^{\infty} a^j v_{t-j}\right] + Var[y_t] = \sigma_v^2 \frac{1}{1-a^2} + \sigma_{uv} \frac{1}{1-a^2} = \frac{\sigma_v^2 + \sigma_{uv}}{1-a^2}
$$

Even though this exercise might seem lengthy, note that all we did is applying the variance/covariance rules and the geometric series.

Exercise 8

a) Yes. Since $ln(S_t)$ depends on the RV W_t, it is stochastic itself.

b) Yes. Monotone transformations of RVs are of course still RVs.

c) No. As noted above, S_t depends on W_t, which is not iid. To see this, note that e.g. $W_1 \sim N(0,1)$ and $W_5 \sim N(0,5)$, so W_1 and W_5 definitely follow different distributions. We don't have any information on whether $W_1, W_2, ..$ are independent of each other.

d)

$$\begin{aligned} ln(S_t) &= ln(S_0) + \mu_S t + \sigma_S W_t \qquad |e^{(\cdot)} \\ S_t &= S_0 e^{\mu_S t + \sigma_S W_t} \qquad |E[(\cdot)] \\ E[S_t] &= S_0 E\left[e^{\mu_S t + \sigma_S W_t}\right] \qquad |\text{lognormal distribution, see (81)} \\ E[S_t] &= S_0 e^{E[\mu_S t + \sigma_S W_t] + \frac{1}{2} Var[\mu_S t + \sigma_S W_t]} \\ E[S_t] &= S_0 e^{\mu_S t + \frac{1}{2}\sigma_S^2 t} \end{aligned}$$

e)

$$\begin{aligned} S_t &= S_0 e^{\mu_S t + \sigma_S W_t} \qquad |Var[(\cdot)] \\ Var[S_t] &= S_0^2 Var\left[e^{\mu_S t + \sigma_S W_t}\right] \qquad |\text{lognormal distribution, see (81)} \\ Var[S_t] &= S_0^2 e^{2E[\mu_S t + \sigma_S W_t] + Var[\mu_S t + \sigma_S W_t]}(e^{Var[\mu_S t + \sigma_S W_t]} - 1) \\ Var[S_t] &= S_0^2 e^{2\mu_S t + \sigma_S^2 t}(e^{\sigma_S^2 t} - 1) \end{aligned}$$

Exercise 9: Right or Wrong?

a) Wrong: Only continuous RVs, for which the integral $\int_{-\infty}^{\infty} f_X(x)dx$ exists have a PDF.

b) Wrong, it's the other way round: Suppose the following holds for the entire population: The population consists of 10 individuals, and for the variable X, $E[X] = 12.5$. Now depending on the sample we draw from this population, we get different values for the sample mean $\bar{x}_n$ where n indicates the number of observations: Possible values might be $\bar{x}_3 = 10$, $\bar{x}_8 = 5.5$, $\bar{x}_{10} = 12.5$, $\bar{x}_3 = 7, \ldots$

c) Right

d) Wrong. The expectation operator is only invariant to linear transformations. E.g. $E[\frac{1}{X}] = \frac{1}{E[X]}$ does not necessarily hold.

e) Right. By definition of 1st centralized moment: $E[X - \mu_X] = E[X - E[X]] = E[X] - E[X] = 0$

f) Wrong. Replace 'exactly' by 'approximately'. As n increases, a normal distribution becomes a better and better approximation, yet the exact distribution usually is painfully complicated and not the normal distribution (unless the underlying RVs follow a normal distribution of course) . We though are usually happy with the approximation made by the CLT, which is very close to the exact distribution if roughly $n > 30$.

Exercise 10

(1) To answer this we take the differences $Glaucoma - Normal$ for each patient and test for the mean of those differences being zero.
For upper-tail test:

$$H_0 : \theta = 0 \quad \text{against} \quad H_1 : \theta > 0.$$

For lower-tail test:

$$H_0 : \theta = 0 \quad \text{against} \quad H_1 : \theta < 0.$$

For two-sides test:

$$H_0 : \theta = 0 \quad \text{against} \quad H_1 : \theta \neq 0.$$

Assuming the data to be normally distributed (for the sake of this example), the mean difference is $\overline{x} = -4$ and the estimated standard deviation is $s = 10.744$ (We can use Stata command *sum difference, detail* to get the mean, std.dev and number of observations as well).
Under H_0 we obtain a t-statistic of

$$t_{data} = \sqrt{n}\frac{\overline{x}}{s} = \frac{-4\sqrt{8}}{10.744} = -1.053.$$

The one-side critical value of t distribution with 7 degree of freedom at the level of 0.05 is 1.8946, and the two-side critical value is 2.3646. Thus the rejection region for upper-tail test is t > 1.8946, for lower-tail test is t < -1.8946, and for two-sides test is |t|> 2.3646.
Since t = - 1.053 doesn't fall into any rejection regions above, so we can reject null hypothesis in neither case. So we can't say that there is a difference in corneal thickness between the eyes.
(2) We calculate the corresponding p-values using stata. The command *ttail(df, t)* reports the the reverse cumulative (upper tail) t distribution and can be used to calculate p-value $= Pr(t > t_{data})$. df is the degrees of freedom and is 7 here; t is the t statistic and here $t_data = -1.053$.
For upper-tail test, the p-value is 0.836 > 0.1. The command is *dis ttail(7, -1.053)*
For lower-tail test, the p-value is 0.163 > 0.1. The command is *dis ttail(7, 1.053)*
For two-sides test, the p-value is 0.327 > 0.1. The command is *dis ttail(7, 1.053)*2*
We cannot reject the null hypothesis of no difference in corneal thickness.

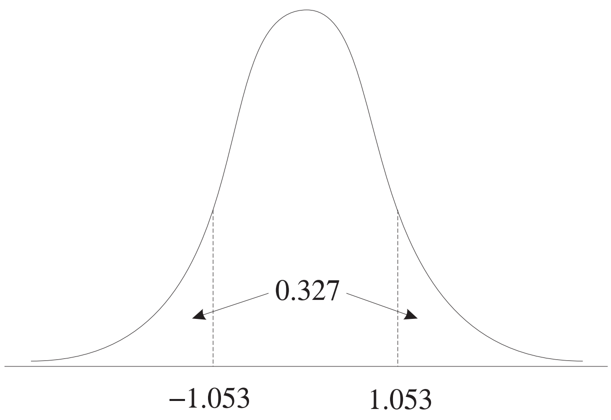

Figure 1 Graph of $t(7)$ p.d.f.

(3) The results are consistent to our calculation in (1) and (2).

```
. ttest difference=0

One-sample t test
------------------------------------------------------------------------------
Variable |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
differ~e |       8          -4    3.798496    10.74377   -12.98202    4.982016
------------------------------------------------------------------------------
    mean = mean(difference)                                       t =  -1.0530
Ho: mean = 0                                     degrees of freedom =        7

    Ha: mean < 0                 Ha: mean != 0                 Ha: mean > 0
 Pr(T < t) = 0.1637         Pr(|T| > |t|) = 0.3273          Pr(T > t) = 0.8363
```