PRINCIPLES OF ECONOMETRICS

5TH EDITION

ANSWERS TO ODD-NUMBERED EXERCISES IN APPENDIX B

EXERCISE B.1

(a) $$E\left(\bar{X}\right)=E\left[\frac{1}{n}\left(X_1+X_2+\ldots+X_n\right)\right]=\frac{1}{n}\left(E(X_1)+E(X_2)+\cdots+E(X_n)\right)$$

$$=\frac{1}{n}\left(\mu+\mu+\cdots+\mu\right)=\frac{n\mu}{n}=\mu$$

(b) $$\operatorname{var}\left(\bar{X}\right)=\operatorname{var}\left(\frac{1}{n}\left(X_1+X_2+\cdots+X_n\right)\right)$$

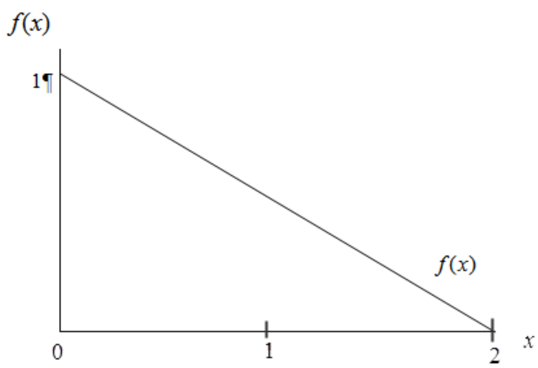
$$=\frac{1}{n^2}\left(\operatorname{var}\left(X_1\right)+\operatorname{var}\left(X_2\right)+\cdots+\operatorname{var}\left(X_n\right)\right)$$

$$=\frac{1}{n^2}n\sigma^2=\frac{\sigma^2}{n}$$

Since $X_1, X_2, \ldots, X_n$ are independent random variables, their covariances are zero. This result was used in the second line of the equation which would contain terms like $\operatorname{cov}(X_i, X_j)$ if these terms were not zero.

EXERCISE B.3

(a) The probability density function is shown below.



(b) Total area of the triangle is half the base multiplied by the height; i.e., the area is $0.5\times2\times1=1$

(c) When $x = 1$, $f(x)=f(1)=\frac{1}{2}$.

Using geometry, $P(X\geq1)$ is given by the area to the right of 1 which is $P(X\geq1)=\frac{1}{2}\times1\times\frac{1}{2}=\frac{1}{4}$.

Using integration,

$$P(X\geq1)=\int_1^2\left(-\tfrac{1}{2}x+1\right)dx=\left(-\tfrac{1}{4}x^2+x\right)\Big|_1^2=(-1+2)-\left(-\tfrac{1}{4}+1\right)=\tfrac{1}{4}$$

(d) When $x = \frac{1}{2}$, $f\left(\frac{1}{2}\right) = \frac{3}{4}$

Using geometry,

$$P\left(X \le \tfrac{1}{2}\right) = 1 - P\left(X > \tfrac{1}{2}\right) = 1 - \tfrac{1}{2} \times 1\tfrac{1}{2} \times \tfrac{3}{4} = \tfrac{7}{16}$$

Using integration,

$$P\left(X \le \tfrac{1}{2}\right) = \int_0^{1/2}\left(-\tfrac{1}{2}x + 1\right)dx = \left(-\tfrac{1}{4}x^2 + x\right)\Big|_0^{1/2} = -\tfrac{1}{16} + \tfrac{1}{2} = \tfrac{7}{16}$$

(e) For a continuous random variable the probability of observing a single point is zero.

Thus, $P\left(X = 1\tfrac{1}{2}\right) = 0$.

(f) The mean is given by

$$E(X) = \int_0^2 x f(x)dx = \int_0^2\left(-\tfrac{1}{2}x^2 + x\right)dx = \left(-\tfrac{1}{6}x^3 + \tfrac{1}{2}x^2\right)\Big|_0^2 = \left(-\tfrac{8}{6} + \tfrac{4}{2}\right) = \tfrac{2}{3}$$

The second moment is given by

$$E(X^2) = \int_0^2 x^2 f(x)dx = \int_0^2\left(-\tfrac{1}{2}x^3 + x^2\right)dx = \left(-\tfrac{1}{8}x^4 + \tfrac{1}{3}x^3\right)\Big|_0^2 = \left(-\tfrac{16}{8} + \tfrac{8}{3}\right) = \tfrac{2}{3}$$

The variance is given by

$$\text{var}(X) = E(X^2) - [E(X)]^2 = \tfrac{2}{3} - \left(\tfrac{2}{3}\right)^2 = \tfrac{2}{9}$$

(g) The cumulative distribution function is given by

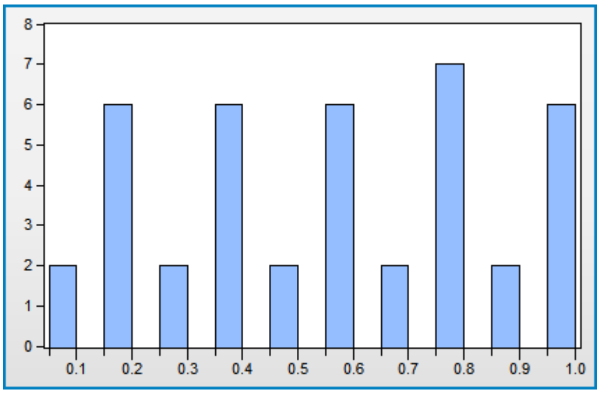
$$F(x) = \int_0^x f(t)dt = \int_0^x\left(-\tfrac{1}{2}t + 1\right)dt = \left(-\tfrac{1}{4}t^2 + t\right)\Big|_0^x = x\left(-\frac{x}{4} + 1\right)$$

EXERCISE B.5

After setting up a workfile for 41 observations, the following EViews program can be used to generate the random numbers

```
series x
x(1)=79
scalar m=100
scalar a=263
scalar cee=71
for !i= 2 to 41
scalar q=a*x(!i-1)+cee
x(!i)=q-m*@ceiling(q/m)+m
next
series u=x/m
```

If the random number generator has worked well, the observations in U should be independent draws of a uniform random variable on the (0,1) interval. A histogram of these numbers follows:



These numbers are far from random. There are no observations in the intervals (0.10,0.15), (0.20,0.25), (0.30,0.35), …. Moreover, the frequency of observations in the intervals (0.05,0.10), (0.25,0.30), (0.45,0.50), … is much less than it is in the intervals (0.15,0.20), (0.35,0.40), (0.55,0.60), … The random number generator is clearly not a good one.

EXERCISE B.7

Let $E_{X,Y}$ be an expectation taken with respect to the joint density for (X,Y); E_X and E_Y are expectations taken with respect to the marginal distributions of X and Y, and $E_{Y|X}$ is an expectation taken with respect to the conditional distribution of Y given X.

Now, $\text{cov}\big(Y, g(X)\big) = 0$ if $E_{X,Y}\big(Y \times g(X)\big) = E_{X,Y}(Y) \times E_{X,Y}\big(g(X)\big)$.

Using iterated expectations, we can write

$$\begin{aligned}
E_{X,Y}\big(Y \times g(X)\big) &= E_X\left[E_{Y|X}\big(Y \times g(X)\big)\right] \\
&= E_X\left[g(X)E_{Y|X}(Y)\right] \\
&= E_X\left[g(X)\right] \times E_Y(Y) \\
&= E_{X,Y}\left[g(X)\right] \times E_{X,Y}(Y)
\end{aligned}$$

EXERCISE B.9

The cumulative distribution function for X is given by $F(x) = \int_0^x \left(\frac{3t^2}{8}\right) dt = \frac{t^3}{8}\Big|_0^x = \frac{x^3}{8}$.

(a) $P\left(0 < X < \frac{1}{2}\right) = F\left(\frac{1}{2}\right) = \frac{1}{64}$

(b) $P(1 < X < 2) = F(2) - F(1) = \frac{7}{8}$

EXERCISE B.11

After setting up a workfile for 41 observations, the following EViews program can be used to generate the random numbers *U1*.

```
series x
x(1)=1234567
scalar m=2^32
scalar a=1103515245
scalar cee=12345
for !i= 2 to 1001
scalar q=a*x(!i-1)+cee
x(!i)=q-m*@ceiling(q/m)+m
next
series u1=x/m
```

If the random number generator has worked well, the observations on *U1* should be independent draws of a uniform random variable on the (0,1) interval. Histograms of these numbers and those from *U2* obtained using the seed value x(1)=95992 follow:

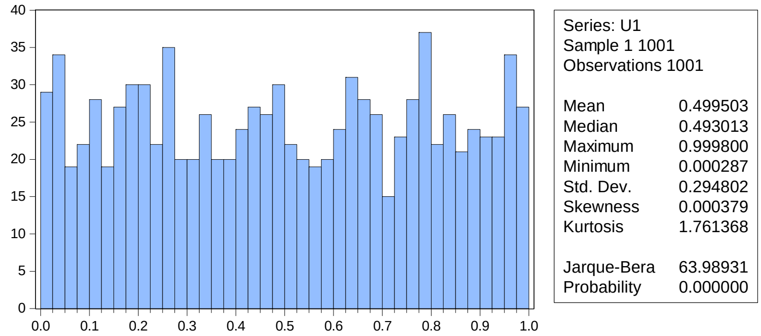


Figure xr-b.11(a) Histogram and summary statistics for *U1*

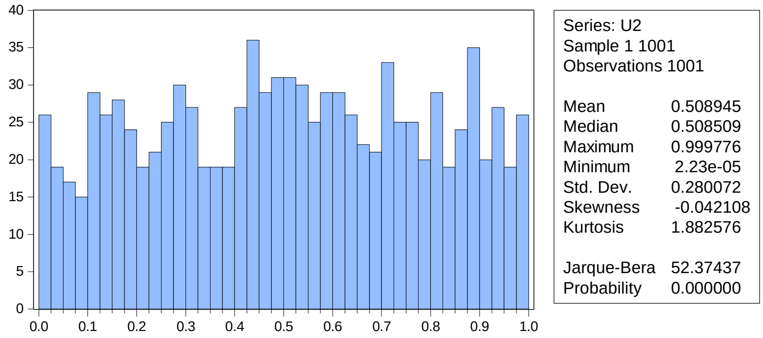


Figure xr-b.11(b) Histogram and summary statistics for *U2*

The histograms are approximately uniformly distributed, implying the random number generator is a good one. The sample means, standard deviations and correlation are

$$U1: \quad \bar{X} = 0.4995 \quad s = 0.2948 \quad \text{cor}(U1, U2) = 0.0471$$

$$U2: \quad \bar{X} = 0.5089 \quad s = 0.2801$$

These sample quantities are very close to the population values $\mu = 0.5$, $\sigma = 0.2887$ and $\rho = 0$.

EXERCISE B.13

(a) The volume under the joint *pdf* is

$$\int_0^2 \int_0^y \left(\frac{1}{2}\right) dx\,dy = \int_0^2 \left[\frac{x}{2}\Big|_0^y \right] dy = \int_0^2 \left(\frac{y}{2}\right) dy = \frac{y^2}{4}\Big|_0^2 = 1$$

(b) The marginal *pdf* for X is $f(x) = 1 - \dfrac{x}{2}$.

The marginal *pdf* for Y is $f(y) = \dfrac{y}{2}$.

(c) $$P\left(X < \tfrac{1}{2}\right) = \frac{7}{16}$$

(d) The *cdf* for Y is $F(y) = \dfrac{y^2}{4}$.

(e) The conditional *pdf* $f(x \mid y)$ is given by

$$f(x \mid y) = \frac{f(x,y)}{f(y)} = \frac{1/2}{y/2} = \frac{1}{y} \qquad \text{implying} \qquad f\left(x \,\middle|\, Y = \tfrac{3}{2}\right) = \frac{2}{3}$$

The required probability is $P\left(X < \tfrac{1}{2} \,\middle|\, Y = \tfrac{3}{2}\right) = \dfrac{1}{3}$.

X and Y are not independent because $P\left(X < \tfrac{1}{2} \,\middle|\, Y = \tfrac{3}{2}\right) \neq P\left(X < \tfrac{1}{2}\right)$.

(f) The mean of Y is $E(Y) = \dfrac{4}{3}$.

The second moment of Y is $E(Y^2) = 2$.

The variance of Y is $\operatorname{var}(Y) = \dfrac{2}{9}$.

(g) From part (e),

$$E(X \mid Y) = \int_0^y x\, f(x \mid y) dx = \int_0^y \left(\frac{x}{y}\right) dx = \frac{x^2}{2y}\Big|_0^y = \frac{y}{2}$$

$$E(X) = E_Y\left[E(X \mid Y)\right] = \int_0^2 \left(\frac{y}{2}\right) f(y) dy = \int_0^2 \left(\frac{y^2}{4}\right) dy = \frac{y^3}{12}\Big|_0^2 = \frac{2}{3}$$

We can check this result by using the marginal *pdf* for X to find $E(X)$:

$$E(X) = \int_0^2 x\, f(x) dx = \int_0^2 \left(x - \frac{x^2}{2}\right) dx = \left(\frac{x^2}{2} - \frac{x^3}{6}\right)\Big|_0^2 = 2 - \frac{4}{3} = \frac{2}{3}$$

EXERCISE B.15

(a) $$\text{var}(X) = E\left(X^2\right) - \left[E(X)\right]^2 = 8.9 - 2.7^2 = 1.61$$

(b) $$\text{var}\left(X \mid Y = 2\right) = E\left(X^2 \mid Y = 2\right) - \left[E\left(X \mid Y = 2\right)\right]^2 = 9.3 - 2.7^2 = 2.01$$

$$\text{var}\left(X \mid Y = 3\right) = E\left(X^2 \mid Y = 3\right) - \left[E\left(X \mid Y = 3\right)\right]^2 = 8.9 - 2.7^2 = 1.61$$

These two conditional variances are not equal.

(c) $$E(X \mid Y = 1) = E\left(X \mid Y = 2\right) = E\left(X \mid Y = 3\right) = 2.7\,.$$

$$E\left(X\right) = \sum_{i=1}^{3} E\left(X \mid Y = i\right) P\left(Y = i\right) = 2.7 \times 0.2 + 2.7 \times 0.5 + 2.7 \times 0.3 = 2.7$$

(d) $$E\left(XY\right) = \sum_{x=1}^{4}\sum_{y=1}^{3} xyf(x, y) = 5.67$$

(e) $$\text{cov}\left(X, Y\right) = E\left(XY\right) - E\left(X\right)E\left(Y\right) = 5.67 - 2.7 \times 2.1 = 0$$

This is an example of random variables whose covariance is zero despite the fact that they are not independent.

(f) The correlation between X and Y is $\text{corr}(X, Y) = \dfrac{\text{cov}(X, Y)}{\sqrt{\text{var}(X)\,\text{var}(Y)}} = 0$

EXERCISE B.17

(a) The values that X can take and their probabilities of occuring, where X is the number of heads occurring in two flips of a fair coin, are given in the first two columns of the following table.

x	$f(x)$	$F(x)$
0	$1/4$	$1/4$
1	$1/2$	$3/4$
2	$1/4$	1

$$P(X \le 1.5) = P(X = 0) + P(X = 1) = \frac{1}{4} + \frac{1}{2} = \frac{3}{4}$$

(b) The values of the cumulative distribution function of X are given in the third column of the table in part (a).

$$P(X \le 1.5) = F(1.5) = F(1) = \frac{3}{4}$$

(c) The *pdf* for $W = 2X$ is given in the following table.

w	$f(w)$
0	$1/4$
2	$1/2$
4	$1/4$

(d) Expected winnings are given by

$$E(W) = 0 \times \frac{1}{4} + 2 \times \frac{1}{2} + 4 \times \frac{1}{4} = 2$$

(e) Let $Y = 1$ if the first flip is a head and $Y = 0$ if the first flip is a tail. Then,

$$P(X = 0 \mid Y = 1) = 0$$

$$P(X = 1 \mid Y = 1) = \frac{1}{2}$$

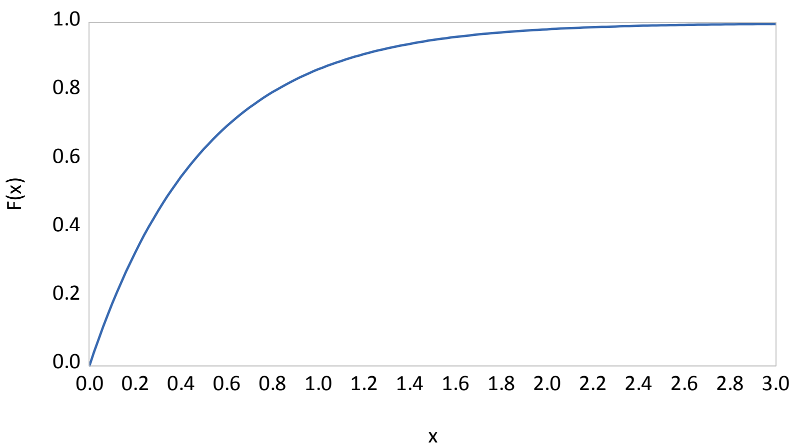
$$P(X = 2 \mid Y = 1) = \frac{1}{2}$$

(f) The conditional expectation of $W = 2X$ given that the first flip is a head is

$$E(W \mid Y = 1) = 2 \times \frac{1}{2} + 4 \times \frac{1}{2} = 3$$

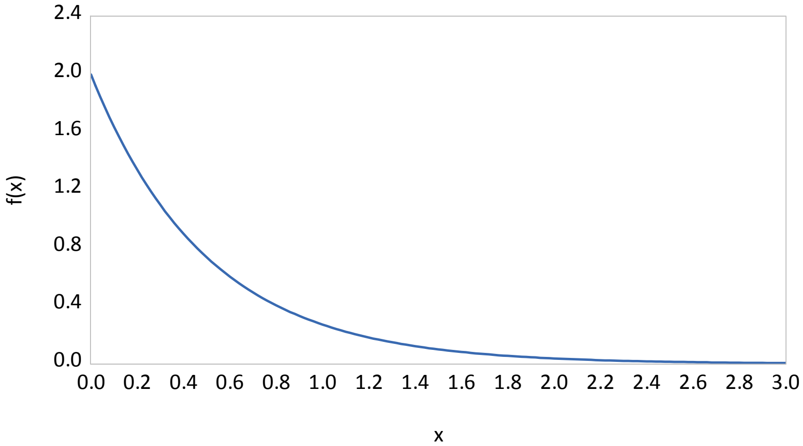
EXERCISE B.19

(a) A graph of the *cdf* follows

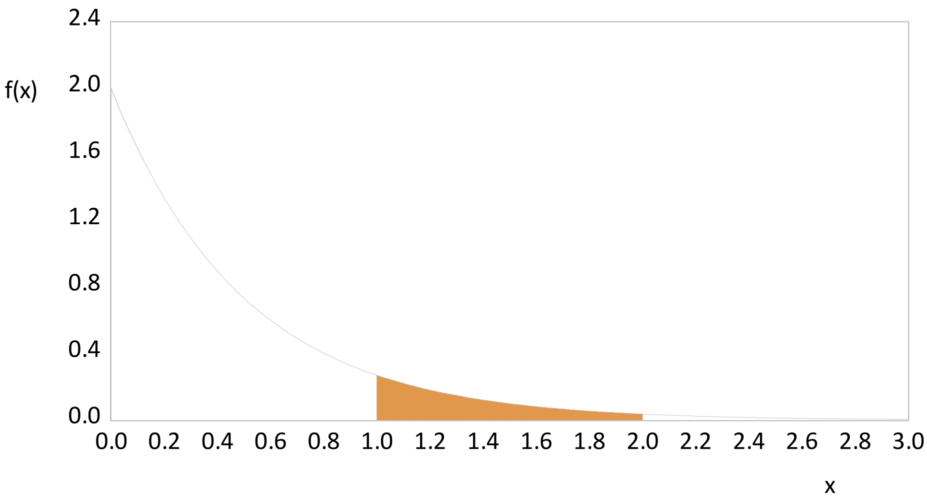


(b) $P(1 < X < 2) = F(2) - F(1) = 0.1353 - 0.0183 = 0.1170$

(c) $f(x) = \dfrac{dF(x)}{dx} = \dfrac{d\left(1 - e^{-2x}\right)}{dx} = 2e^{-2x}$



(d) $P(1 < X < 2)$ is given by the shaded area in the following graph.



EXERCISE B.21

The marginal *pdf*'s for Y and X are

y	$f(y)$
1	0.2
2	0.5
3	0.3

x	$f(x)$
1	0.3
2	0.1
3	0.2

(a) $\text{var}(Y) = 0.49$

(b) The conditional distributions for Y for each possible value of X are given in the following table.

y	$f(y\mid X=1)$	$f(y\mid X=2)$	$f(y\mid X=3)$	$f(y\mid X=4)$
1	$\frac{0.01}{0.3}=0.0333$	$\frac{0.07}{0.1}=0.7$	$\frac{0.09}{0.2}=0.45$	$\frac{0.03}{0.4}=0.075$
2	$\frac{0.2}{0.3}=0.6667$	$\frac{0}{0.1}=0$	$\frac{0.05}{0.2}=0.25$	$\frac{0.25}{0.4}=0.625$
3	$\frac{0.09}{0.3}=0.3$	$\frac{0.03}{0.1}=0.3$	$\frac{0.06}{0.2}=0.3$	$\frac{0.12}{0.4}=0.3$

$E(Y\mid X=1)=2.26667 \qquad E(Y\mid X=2)=1.6 \qquad E(Y\mid X=3)=1.85 \qquad E(Y\mid X=4)=2.225$

(c) We have,

$$\sum_{x=1}^{4}\left[E\left(Y\mid X=x\right)-E\left(Y\right)\right]^2 f\left(x\right)=\left(2.26667-2.1\right)^2\times 0.3+\left(1.6-2.1\right)^2\times 0.1$$
$$+\left(1.85-2.1\right)^2\times 0.2+\left(2.225-2.1\right)^2\times 0.4$$
$$=0.0521$$

This represents the first term in (B.27). It is the variance of the conditional mean.

(d)
$$\text{var}(Y\mid X=1)=E\left(Y^2\mid X=1\right)-\left[E(Y\mid X=1)\right]^2=5.4-2.26667^2=0.2622$$
$$\text{var}(Y\mid X=2)=E\left(Y^2\mid X=2\right)-\left[E(Y\mid X=2)\right]^2=3.4-1.6^2=0.84$$
$$\text{var}(Y\mid X=3)=E\left(Y^2\mid X=3\right)-\left[E(Y\mid X=3)\right]^2=4.15-1.85^2=0.7275$$
$$\text{var}(Y\mid X=4)=E\left(Y^2\mid X=4\right)-\left[E(Y\mid X=4)\right]^2=5.275-2.225^2=0.3244$$

(e)
$$\sum_{x=1}^{4}\left[\text{var}\left(Y\mid X=x\right)\right]f\left(x\right)=0.2622\times 0.3+0.84\times 0.1+0.7275\times 0.2+0.3244\times 0.4=0.4379$$

This represents the second term in (B.27). It is the mean of the conditional variances.

(f) We have,

$$\text{var}(Y)=\sum_{x=1}^{4}\left[E\left(Y\mid X=x\right)-E\left(Y\right)\right]^2 f\left(x\right)+\sum_{x=1}^{4}\left[\text{var}\left(Y\mid X=x\right)\right]f\left(x\right)$$
$$=0.0521+0.4379=0.49$$

This result agrees with that from part (a).

EXERCISE B.23

(a) $P(X=2\mid \mu=2)=0.27067$

(b) $P(X\geq 2)=1-P(X=0)-P(X=1)=0.59399$

(c) $P(X\geq 2\mid X\geq 1)=\dfrac{P\left[(X\geq 2)\cap(X\geq 1\right]}{P(X\geq 1)}=\dfrac{0.59399}{1-0.13534}=0.68696$

EXERCISE B.25

(a) Chebyshev's inequality is $P\left(|X-\mu|\geq\varepsilon\right)\leq\sigma^2/\varepsilon^2$. When $\varepsilon=k\sigma$, it becomes

$$P\left(|X-\mu|\geq k\sigma\right)\leq\frac{1}{k^2}$$

(b) If $k=2$, $\mu=1$ and $\sigma^2=1$, then Chebyshev's inequality becomes $P\left(|X-1|\geq 2\right)\leq 0.25$.

For a normal distribution where $X\sim N(1,1)$ and $Z=(X-1)\sim N(0,1)$,

$$P\left(|X-1|\geq 2\right)=2\times\left(1-P(Z\leq 2)\right)=2\times\left(1-0.9772\right)=0.0456$$

Since $0.0456<0.25$, this result is in line with Chebyshev's inequality.

(c) If U is a uniform random variable on the interval [0, 1], then $\mu=E(U)=0.5$, and $\sigma^2=\text{var}(U)=1/12$. Thus,

$$P\left(|U-0.5|\geq 2\sigma\right)=2\times P\left(U-0.5\geq\frac{2}{\sqrt{12}}\right)=2\times P\left(U\geq 1.077\right)=0$$

Since $0<0.25$, this result is in line with Chebyshev's inequality.

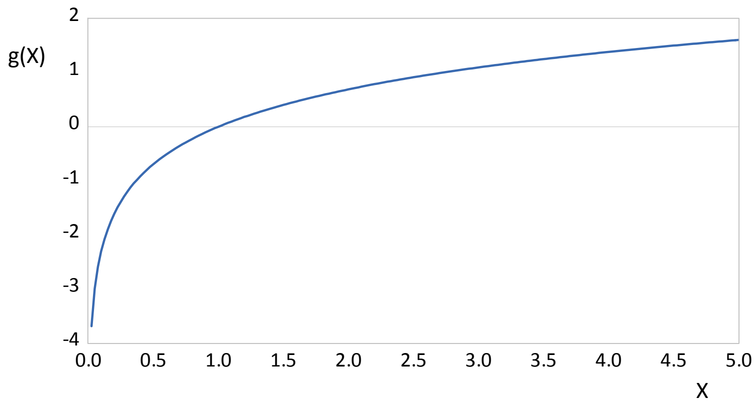
(d) If Y is a binomial random variable with n = 10 and p = 0.8, then $\mu=E(Y)=np=8$, and $\sigma=\sqrt{\text{var}(Y)}=\sqrt{np(1-p)}=1.26491$. Thus,

$$\begin{aligned}P\left(|Y-\mu|\geq 2\sigma\right)&=P\left(Y-\mu\geq 2\sigma\right)+P\left(\mu-Y\geq 2\sigma\right)\\&=P\left(Y\geq 8+2\times 1.26491\right)+P\left(Y\leq 8-2\times 1.26491\right)\\&=P\left(Y\geq 10.5298\right)+P\left(Y\leq 5.4702\right)=0+P(Y\leq 5)=0.03279\end{aligned}$$

Since $0.03279<0.25$, this result is in line with Chebyshev's inequality.

EXERCISE B.27

(a) If $g(X)=\ln(X)$, then $\dfrac{d^2g}{dX^2}=-\dfrac{1}{X^2}$. Because $\dfrac{d^2g}{dX^2}<0$ for $X>0$, $g(X)=\ln(X)$ is concave for $X>0$.



(b) $E(X)=3$ and $E\left[\ln(X)\right]=1.0227$.

Since $\ln(3)=1.0986>1.0227$, the inequality $\ln\left[E\left(X\right)\right]\geq E\left[\ln\left(X\right)\right]$ is satisfied.

(c) $\ln\left(\sum_{i=1}^{4}x_i\Big/4\right)=\ln(2.5)=0.9163$ and $\sum_{i=1}^{4}\ln\left(x_i\right)\Big/4=0.7945$.

Hence, $\ln\left(\sum_{i=1}^{4}x_i\Big/4\right)\geq\sum_{i=1}^{4}\ln\left(x_i\right)\Big/4$.

EXERCISE B.29

(a) $P\left(X^2\geq 5\right)=P\left(X\geq 3\right)=0.7$ and $E\left(X^2\right)=10$.

Thus, $P\left(X^2\geq 5\right)=0.7\leq E\left(X^2\right)\Big/5=2$.

(b) When $g\left(X\right)=\left(X-\mu_X\right)^2$ and $c=k^2\sigma_X^2$, Markov's Inequality $P\left[g\left(X\right)\geq c\right]\leq c^{-1}E\left[g\left(X\right)\right]$ becomes

$$P\left[\left(X-\mu_X\right)^2\geq k^2\sigma_X^2\right]\leq\frac{\sigma_X^2}{k^2\sigma_X^2}$$

from which we obtain

$$P\left[\left|X-\mu_X\right|\geq k\sigma_X\right]\leq\frac{1}{k^2}$$

Let $\varepsilon=k\sigma_X$ so that $1\big/k^2=\sigma_X^2\big/\varepsilon^2$, then we get Chebyshev's Inequality,

$$P\left[\left|X-\mu_X\right|\geq\varepsilon\right]\leq\frac{\sigma_X^2}{\varepsilon^2}$$

EXERCISE B.31

(a)
$$E(Y\mid X)=1\times x+0\times(1-x)=x$$

$$E(Y)=E_X\left[E(Y\mid X)\right]=\int_0^1 x\,f(x)dx=\left.\frac{x^2}{2}\right]_0^1=\frac{1}{2}$$

(b)
$$\operatorname{var}(Y\mid X)=E\left(Y^2\mid X\right)-\left[E(Y\mid X)\right]^2=1^2\times x+0^2\times(1-x)-x^2=x-x^2$$

$$\operatorname{var}(Y)=E_X\left[\operatorname{var}(Y\mid X)\right]+\operatorname{var}_X\left[E(Y\mid X)\right]=\int_0^1\left(x-x^2\right)dx+\operatorname{var}_X(x)$$

Because X is a uniform random variable over the interval (0,1), $\operatorname{var}_X(x)=\dfrac{1}{12}$. Thus,

$$\operatorname{var}(Y)=\left.\left(\frac{x^2}{2}-\frac{x^3}{3}\right)\right|_0^1+\frac{1}{12}=\frac{1}{2}-\frac{1}{3}+\frac{1}{12}=\frac{1}{4}$$