For Instructors

Solutions to End-of-Chapter Exercises

Chapter 2  
Review of Probability

2.1. (a) Probability distribution function for *Y*

|  |  |  |  |
| --- | --- | --- | --- |
| Outcome (number of heads) | *Y*  0 | *Y*  1 | *Y*  2 |
| Probability | 0.25 | 0.50 | 0.25 |

(b) Cumulative probability distribution function for *Y*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Outcome (number of heads) | *Y*  0 | 0 ≤ *Y*  1 | 1 ≤ Y  2 | *Y* ≥ 2 |
| Probability | 0 | 0.25 | 0.75 | 1.0 |

(c) .

Using Key Concept 2.3: 

and



so that



2.2. We know from Table 2.2 that     So

(a) 

(b) 

(c) 



2.3. For the two new random variables  and  we have:

(a) 

(b) 

(c) 



2.4. (a) 

(b) 

(c) , and var(*X*) = *E*(*X*2)−[*E*(*X*)]2 = 0.3 −0.09 = 0.21. Thus *σ* = = 0.46.

 To compute the skewness, use the formula from exercise 2.21:



Alternatively, 

Thus, skewness 

To compute the kurtosis, use the formula from exercise 2.21:



Alternatively, 

Thus, kurtosis is 

2.5. Let *X* denote temperature in °F and *Y* denote temperature in °C. Recall that *Y*  0 when *X*  32 and *Y* 100 when *X*  212; this implies  Using Key Concept 2.3, *μX*  70oF implies that  and *σX*  7oF implies 

2.6. The table shows that        

(a) 

(b) 

(c) Calculate the conditional probabilities first:



The conditional expectations are



(d) Use the solution to part (b),

Unemployment rate for college graduates  1  *E*(*Y*|*X*  1)  1  0.974  0.026

Unemployment rate for non-college graduates  1  *E*(*Y*|*X*  0)  1  0.944  0.056

(e) The probability that a randomly selected worker who is reported being unemployed is a college graduate is



The probability that this worker is a non-college graduate is



(f) Educational achievement and employment status are not independent because they do not satisfy that, for all values of *x* and *y*,



For example, from part (e)  while from the table Pr(*X*  0)  0.659.

2.7. Using obvious notation,  thus  and  This implies

(a)  per year.

(b)  , so that  Thus   where the units are squared thousands of dollars per year.

(c)  so that  and  thousand dollars per year.

(d) First you need to look up the current Euro/dollar exchange rate in the Wall Street Journal, the Federal Reserve web page, or other financial data outlet. Suppose that this exchange rate is *e* (say *e*  0.80 Euros per dollar); each 1 dollar is therefore with *e* Euros. The mean is therefore  
*e* × *μC* (in units of thousands of Euros per year), and the standard deviation is *e* × *σC* (in units of thousands of Euros per year). The correlation is unit-free, and is unchanged.

2.8.   With 



|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 2.9. |  | | **Value of *Y*** | | | | | **Probability Distribution of *X*** |
| **14** | **22** | **30** | **40** | **65** |
|  | **Value of *X*** | 1 | 0.02 | 0.05 | 0.10 | 0.03 | 0.01 | 0.21 |
| 5 | 0.17 | 0.15 | 0.05 | 0.02 | 0.01 | 0.40 |
| 8 | 0.02 | 0.03 | 0.15 | 0.10 | 0.09 | 0.39 |
|  | **Probability distribution of *Y*** | | 0.21 | 0.23 | 0.30 | 0.15 | 0.11 | 1.00 |

(a) The probability distribution is given in the table above.



(b) The conditional probability of *Y*|*X*  8 is given in the table below

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Value of *Y*** | | | | |
| 14 | 22 | 30 | 40 | 65 |
| 0.02/0.39 | 0.03/0.39 | 0.15/0.39 | 0.10/0.39 | 0.09/0.39 |





(c) 





2.10. Using the fact that if  then  and Appendix Table 1, we have

(a) 

(b) 

(c) 

(d) 

2.11. (a) 0.90

(b) 0.05

(c) 0.05

(d) When  then 

(e)  where thus 

2.12. (a) 0.05

(b) 0.950

(c) 0.953

(d) The *tdf* distribution and *N*(0, 1) are approximately the same when *df* is large.

(e) 0.10

(f) 0.01

2.13. (a) 

(b) *Y* and *W* are symmetric around 0, thus skewness is equal to 0; because their mean is zero, this means that the third moment is zero.

(c) The kurtosis of the normal is 3, so  ; solving yields  a similar calculation yields the results for *W*.

(d) First, condition on  so that 



Similarly,



From the law of iterated expectations



(e)  thus from part (d). Thus skewness  0. Similarly,  and  Thus,

2.14. The central limit theorem suggests that when the sample size (*n*) is large, the distribution of the sample average  is approximately  with  Given  

(a)   and



(b)   and



(c)   and



2.15. (a) 

where *Z* ~ *N*(0, 1). Thus,

(i) *n*  20; 

(ii) *n*  100; 

(iii) *n*  1000; 

(b) 

As *n* get large  gets large, and the probability converges to 1.

(c) This follows from (b) and the definition of convergence in probability given in Key Concept 2.6.

2.16. There are several ways to do this. Here is one way. Generate *n* draws of *Y*, *Y*1, *Y*2, … *Yn*. Let *Xi*  1 if *Yi*  3.6, otherwise set *Xi*  0. Notice that *Xi* is a Bernoulli random variables with *μX*  Pr(*X*  1)  Pr(*Y*  3.6). Compute  Because  converges in probability to *μX*  Pr(*X*  1)  Pr(*Y*  3.6),  will be an accurate approximation if *n* is large.

2.17. *μY* = 0.4 and 

(a) (i) *P*( ≥ 0.43)  

(ii) *P*( ≤ 0.37)  

(b) We know Pr(1.96 ≤ *Z* ≤ 1.96)  0.95, thus we want *n* to satisfy  and  Solving these inequalities yields *n* ≥ 9220.

2.18. 

(a) The mean of *Y* is



The variance of *Y* is



so the standard deviation of *Y* is 

(b) (i)  

(ii) Using the central limit theorem,



2.19. (a) 

(b) 

(c) When  and  are independent,



so





2.20. (a) 

(b) 

where the first line in the definition of the mean, the second uses (a), the third is a rearrangement, and the final line uses the definition of the conditional expectation.

2.21. (a) 

(b) 

2.22. The mean and variance of *R* are given by



where  follows from the definition of the correlation between *Rs* and *Rb*.

(a) 

(b) 

(c) *w*  1 maximizes  for this value of *w*.

(d) The derivative of *σ*2 with respect to *w* is



Solving for *w* yields (Notice that the second derivative is positive, so that this is the global minimum.) With 

2.23. *X* and *Z* are two independently distributed standard normal random variables, so



(a) Because of the independence between  and   and  Thus 

(b)  and 

(c)  Using the fact that the odd moments of a standard normal random variable are all zero, we have  Using the independence between  and  we have  Thus 

(d) 

2.24. (a)  and the result follows directly.

(b) (*Yi*/*σ*) is distributed i.i.d. *N*(0,1),  and the result follows from the definition of a  random variable.

(c) 

(d) Write



which follows from dividing the numerator and denominator by *σ*. *Y*1/*σ* ~ *N*(0,1), ~ , and *Y*1/*σ* and  are independent. The result then follows from the definition of the *t* distribution.

2.25. (a) 

(b) 

(c) 

(d) 

2.26. (a) corr(*Yi*,*Yj*)  , where the first equality uses the definition of correlation, the second uses the fact that *Yi* and *Yj* have the same variance (and standard deviation), the third equality uses the definition of standard deviation, and the fourth uses the correlation given in the problem. Solving for cov(*Yi*, *Yj*) from the last equality gives the desired result.

(b) , so that *E*()  

(c) , so that 



where the fourth line uses  for any variable *a*.

(d) When *n* is large  and , and the result follows from (c).

2.27 (a) *E*(*W*)  *E*[*E*(*W*|*Z*)]  *E*[*E*(*X *)|*Z*]  *E*[*E*(*X*|*Z*)  *E*(*X*|*Z*)]  0.

(b) *E*(*WZ*)  *E*[*E*(*WZ*|*Z*)]  *E*[*ZE*(*W*)|*Z*]  *E*[ *Z* ×0]  0

(c) Using the hint: *V*  *W*  *h*(*Z*), so that *E*(*V*2)  *E*(*W*2)  *E*[*h*(*Z*)2]  2 × *E*[*W* × *h*(*Z*)]. Using an argument like that in (b), *E*[*W* × *h*(*Z*)]  0. Thus, *E*(*V*2)  *E*(*W*2)  *E*[*h*(*Z*)2], and the result follows by recognizing that *E*[*h*(*Z*)2] ≥ 0 because *h*(*z*)2≤ 0 for any value of *z*.