Andrew Shaw

Reading: Stock & Watson 2.1-2.3

1. Describe probability in your own words (note do not use the words from the book, think about it and come up with your own description).

A way to quantify the chance of some event happening found by taking the number of possible times of a desirable outcome and dividing out the total of all possible outcomes.

1. What is the difference between a discrete and continuous variable? Give examples of both types of variables.

Discrete variables represent something separate and distinct. A continuous variable represents something that is infinitely divisible.

Discrete could be number of cars produced.

Continuous could be tire pressure or temperature.

1. Define the cumulative probability distribution function of the following discrete variable. Pr(X=0)=0.05, Pr(X=1)=0.15, Pr(X=2)=0.3, Pr(X=3)=0.5.

{0, for x < 0

{0.05, for 0 ≤ x < 1,

{0.05 + 0.15 = 0.20 , 1 ≤ x < 2,

{0.20 + 0.30 = 0.50, for 2 ≤ x < 3,

{0.50 + 0.50 = 1.00, for x ≥ 3

1. DD brings $20 to school on Mondays, $50 to school on Tuesdays, $35 on Wednesdays, $40 on Thursdays, and $100 on Fridays. If a thief picks a random day of the week to pick DD’s pocket, what is his expected value of their take?

E[X] = .2(20+50+35+40+100) = $49

1. The mean absolute deviation of a set of numbers is the average of all absolute differences each element is from the mean of a set. For example the mean absolute deviation of A{1,1,10} is 4 or

In addition the mean absolute deviation of B{-2,4,10} is 4 or

What is the standard deviation of A and B? Why is the standard deviation a more desirable sample dispersion measure over mean absolute deviation?

A = 4.243

B = 4.899

Standard deviation is preferred because it has properties that work well with normally distributed data, it can be differentiated, squaring helps highlight outliers too.

Table 1

|  |  |  |  |
| --- | --- | --- | --- |
|  | <100 hours of study (X=0) | >100 hours of study  (X=1) |  |
| Grade <A (Y=0) | a. 0.2 | c. 0.2 | g. 0.4 |
| Grade=A (Y=1) | b. 0.1 | d. 0.5 | h. 0.6 |
|  | e. 0.3 | f. 0.7 | 1.00 |

1. a, b, c, d in table 1 represent the joint distribution of the probability a student gets an A in my class given they study < 100 or greater than 100 hours. e,f,g,h represent the marginal distribution. Fill in c,d,e, and f.
2. In table 1, what is the conditional probability students get an A in my class if they study <100 hours and >100 hours? Show your formula.

Probability of studying less than 100 hrs:

Pr(Y = 1 | X = 0) = 0.1/0.3 = 0.3333

Probability of studying more than 100 hrs:

Pr(Y = 1 | X = 1) = 0.5/0.7 = 0.7143

Coin toss Game: The amount you take depends on the number of times you get heads in a row, and the coin will be flipped 4 times. The game ends when tails is flipped. The probability you get heads on any given toss is 0.5. It costs $1 to play and you get $1 back if you get heads once, $3 back if you get heads twice, $7 back if you get heads three times and $18 if you get heads four times in a row.

1. Is this a game you should play? (Hint use the conditional expectation and law of iterated expectation formulas.)

Yes, you should play. Net payout is about $1.19

E(Y|X = x) = 0.5\*0 (T) + 0.25\*1 (HT) + 0.125\*3 (HHT) + 0.0625\*7 (HHHT) + 0.0625\*18 (HHHH) = 2.1875 – 1 = $1.1875

1. What are two real world independent variables?

The chance of rain tomorrow in NYC and the result of a coin flip.

1. Describe the covariance in your own words. (again do not simply reiterate the words of the book). Why is covariance an awkward concept? How does correlation improve on this?

Covariance describes how two variables move in reference to each other (i.e. whether they move in the same direction (positive) or in opposite directions (negative)). The tricky thing about cov is how to interpret because they are in units of the product of the two variables.

Correlation removes this by dividing by the standard deviations of the two variables making correlation unitless.

1. What is so normal about the normal distribution anyway?

It is symmetric about the mean with a standard deviation of 1.

1. I really like normal variables with a mean of zero and a variance of one, how can I change any normal variable into a Z normal variable.

By subtracting the mean from the variable and dividing by its standard deviation to normalize.

1. TT scores 3.63 standard deviations above the mean on a fitness test that is normally distributed. How many people in the world will score better than TT?

About 0.014% of people would score better than TT.

(1-.99986) = 0.00014

1. How do student’s t, chi-square and normal distributions relate to one another? How about the F-distribution?’

The chi-square builds on normal distribution and the t-distribution builds both of those and typically used when standard deviation is unknown and small sample size.

The F distribution is a ratio between two chi-squared independent variables

1. What does it mean that the variance of the sampling distribution of the mean is zero as n goes to infinity? Does it mean that large samples have zero variance?

No, there is still variance it just means with more observations numbers above and below the mean tend to cancel each other out resulting a steadier average.

1. Why is the normal distribution so special?

It’s representative of the Central limit theorem and only requires mean and variance to describe it.