Hickman Problem Set 7

Keith Hickman

October 12, 2017

Problem 1 - Trosset 8.6.4

(a) Approximate the probability that you will make a profit on your investment if you purchase a share of the value stock.

Here, we need to find the probability that after 400 days, our initial value stock investment of 50 will be worth at least 50. I want to know the sum of the iid random variable.

With EX_i of .01 and Var of .01, to calculate the standard deviation, we can take $\sqrt{.01} = .1$. We can estimate the probability of retaining the initial value at $n = 400$ days with `pnorm`. Adding the initial value of 50 to .01 and plugging in $.1 * sqrt(n)$. We get the following equation:

```
1 - pnorm(50,50.01,.1*sqrt(400))

## [1] 0.5019947
```

or ~50.2% probability that the investment will be at least $50 after 400 days.

(b) Approximate the probability that you will make a profit on your investment if you purchase a share of the growth stock.

```
1 - pnorm(50,50.01,.5*sqrt(400))

## [1] 0.5003989
```

Here, because the variance (volatility) is greater, the probability that the initial investment will be at least $50 is lower, or 50.04%.

(c) Approximate the probability that you will make a profit of at least $20 if you purchase a share of the value stock.

```
1 - pnorm(70,50.01,.1*sqrt(400))

## [1] 0
```

The probability is 0 that I would make a profit of $20.

(d) Approximate the probability that you will make a profit of at least $20 if you purchase a share of the growth stock.

```
1 - pnorm(70,50.01,.5*sqrt(400))

## [1] 0.02280418
```

There is a 2% probability that I would make a $20 profit with the growth stock after 400 days, which would represent a 40% return.

(e) Assuming that the growth stock fluctuations and the value stock fluctuations are independent, approximate the probability that, after 400 days, the price of the growth stock will exceed the price of the value stock.

Here, we are interested in the difference between two iid random variables G and V (growth and value). The first step is to calculate the difference between the V and G variables, which can be done by subtracting the $E_G - E_V = .01 - .01 = 0$. The sum of the Variance of V and $G = .25 + .01 = .26$. We now have a new iid random variable Y with mean 0 and variance .26. The standard deviation of $Y = \sqrt{.26} = .509$. To find the probability that Y is greater than 0 after 400 days, we need to find the sum of Y with n = 400. We add the $\mu * n = 400 * 0 = 0 *$ and $Y_\sigma = .509 * \sqrt{400} = 10.19$. We can then use `1 - pnorm` to find the probability that Y will be greater than zero:

```
1-pnorm(0,0,10.19)

## [1] 0.5
```

Thus, there is a 50% chance that the Value stock will be worth more than the Growth stock after 400 days. *This doesn't quite feel right - there's more volatility in the growth stock, so shouldn't there be at least a slightly higher chance that the value stock will be worth more?*

Question 2

Problem 2 - Trosset 8.6.4

(a) Find EX Here, EX = 0 or $(-2 * 3) + (-1 * .6) + (12 * .1) = 0$

(b) Find Var(X). Var X = 16.2

(c) What is the expected value of $\overline{X}$? Here, $E\overline{X}$ = EX = 0

(d) What is the variance of $\overline{X}$ (Note: This will depend on n.) Var $\overline{X}$ = $\frac{16.2}{\sqrt{(n)}}$.

(e) Suppose n = 100. Use the R function `pnorm()` to find the approximate probability that $\overline{X}$ is greater than 0.5

To find the variance of $\overline{X}$, we divide the variance of X by the square root of n, in this case n = 100, $\sqrt{100} = 10$: $16.2/10 = 1.62$

```
1 - pnorm(.5, 0, 1.62)

## [1] 0.3787969
```

Thus, there is a ~37.8% that $\overline{X}$ is greater than .5

Problem 3

I read the data into a variable x using the following command:

```
x = rep(c(1, 2, 3, 4, 5, 6, 7), times = c(27, 34, 16, 13, 6, 3, 1))
x

##   [1] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2
2
##  [36] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 3 3 3 3 3 3 3
3
##  [71] 3 3 3 3 3 3 3 4 4 4 4 4 4 4 4 4 4 4 4 4 5 5 5 5 5 5 6 6 6 7
```

(a) What is the sample mean of our data?

The expected value of the sample mean is equal to the mean of the population.

```
mean(x)

## [1] 2.5
```

(b) What is the estimate of standard deviation of our data.

The standard deviation of a sample mean is equal to the population standard deviation divided by the square root of n.

```
y = sd(x)/sqrt(length(x))
y

## [1] 0.1410638
```

(c) What is the estimated standard error of the sample mean? (That is, based on our answer to (b), what is our estimate for the standard deviation of the distribution of the sample mean?)

The standard error of our sample mean is the standard deviation divided by the square root of n.

```
stderror.x = y/sqrt(length(x))
stderror.x

## [1] 0.01410638
```

(d) Our error is the difference between the sample mean and the population mean. Using the normal distribution, find the approximate probability that the absolute value of the error in a survey of this form and size is less than 0.5.

The normal distribution has a mean μ 0 and standard deviation σ of 1. We need the probability that our error is greater than -.5 and less than .5. Converting to standardized numbers, we want to solve for $P(\frac{-.5}{.1} < \frac{X_{100}}{.1} < \frac{.5}{.1})$ = $P(-5 < Z < 5)$ =

```r
pnorm(5) - pnorm(-5)
```

```
## [1] 0.9999994
```

Or greater than 99.99% certainty that our error will be less than .5.

(e) Can we be reasonably sure that the average household size for all U.S. households is between 2 and 3? (Hint: Use results from parts (a) and (d).)

Yes, we can be reasonably sure that the mean household size is between 2 and 3. According to (a), the mean size is 2.5. According to (d), there is a 99.99% certainty that our error size is less than .5, which when added to or subtracted from the mean, would result in most of the values of our distribution falling between 2 and 3.