605-HW09-CLT-and-Generating-Functions

Michael Y.

October 27, 2019

Contents

HW9 - Sums of RVs; Law of Large Numbers

1. #11 page 363

The price of one share of stock in the Pilsdorff Beer Company (see Exercise 8.2.12) is given by Y_n on the n^{th} day of the year.

Finn observes that the differences $X_n = Y_{n+1} - Y_n$ appear to be independent random variables with a common distribution having mean $\mu = 0$ and variance $\sigma^2 = \frac{1}{4}$.

So, the first difference is $X_1 = Y_2 - Y_1$.
Thus, the price at time $Y_2 = X_1 + Y_1 = X_1 + 100$.
The next difference is $X_2 = Y_3 - Y_2$.
Likewise, the price at time $Y_3 = X_2 + Y_2 = X_2 + X_1 + Y_1 = X_2 + X_1 + 100$.

Continuing, the price at time $Y_{j+1} = X_j + X_{j-1} + X_{j-2} + ... + X_1 + Y_1 = Y_1 + \sum_{i=1}^{j} X_i = 100 + \sum_{i=1}^{j} X_i$, and

the price at $Y_{n+1} = X_n + X_{n-1} + X_{n-2} + ... + X_1 + Y_1 = Y_1 + \sum_{i=1}^{n} X_i = 100 + \sum_{i=1}^{n} X_i$.

This can also be written as $Y_{n+1} - Y_1 = Y_{n+1} - 100 = \sum_{i=1}^{n} X_i$.

Therefore, the price at $Y_{365} = 100 + \sum_{i=1}^{364} X_i$, where each X_i is distributed as $X \sim UNKNOWN] \left(\mu = 0, \sigma^2 = \frac{1}{4}\right)$ so the standard deviation of the individual X_i is $\sigma = \frac{1}{2}$.

Because the average of the X_j is $E[X] = \mu_X = 0$, the expected value of the Y_j is $E[Y] = 100 + E[X] = 100$.

Although we do not know what kind of distribution is follwed by the individual daily price moves X_i , the Central Limit Theorem tells us that the standard deviation of the difference between their average $\frac{1}{n}\sum_{i=1}^{n} X_i$ and its limit (here, $\mu_X = 0$) is distributed as $N(0, \frac{\sigma^2}{n})$.

Thus, $\frac{1}{n}\sum_{i=1}^{n} X_i \sim N(0, \frac{\sigma^2}{n})$;
$\frac{\sqrt{n}}{n}\sum_{i=1}^{n} X_i \sim N(0, \sigma^2)$; and $Y_{n+1} - Y_1 = \sum_{i=1}^{n} X_i \sim N(0, \sigma^2\sqrt{n})$, where $n = 364$.

This can also be written as $Y_{n+1} = Y_1 + \sum_{i=1}^{n} X_i \sim N(Y_1, \sigma^2 n) \sim N\left(100, \frac{n}{4}\right)$.

Here, $Y_{365} \sim N\left(100, \frac{364}{4}\right) = N(100, 91)$.

When using the R function *pnorm* we have to pass in the standard deviation rather than the variance, so below we use $\sqrt{91}$:

If $Y_1 = 100$, estimate the probability that Y_{365} is:

(a) ≥ 100.

```
probge100 = pnorm(100, mean=100, sd=sqrt(91),lower.tail=F)
probge100
```

```
## [1] 0.5
```

$Pr(Y_{365} \geq 100) = 0.5$.

(b) ≥ 110.

```r
probge110 = pnorm(110, mean=100, sd=sqrt(91),lower.tail=F)
probge110
```

```
## [1] 0.147254
```

$Pr(Y_{365} \geq 110) = 0.147254$.

(c) ≥ 120.

```r
probge120 = pnorm(120, mean=100, sd=sqrt(91),lower.tail=F)
probge120
```

```
## [1] 0.0180158
```

$Pr(Y_{365} \geq 120) = 0.018016$.

Simulation, to confirm theory

```r
nsims=1000000
Y365 = rep(NA,nsims)
Y1 = 100
sigma = 0.5
mu = 0
for(i in 1:nsims) {
  normrands=rnorm(364,mu,sigma)
  Y365[i]=Y1+sum(normrands)
}

p100=sum(Y365>=100)
p110=sum(Y365>=110)
p120=sum(Y365>=120)

count=c(p100,p110,p120)
simprob=count/nsims
theory=c(probge100,probge110,probge120)

diff=simprob-theory

result=cbind(cut=c(100,110,120),count,simprob,theory,diff)
result
```

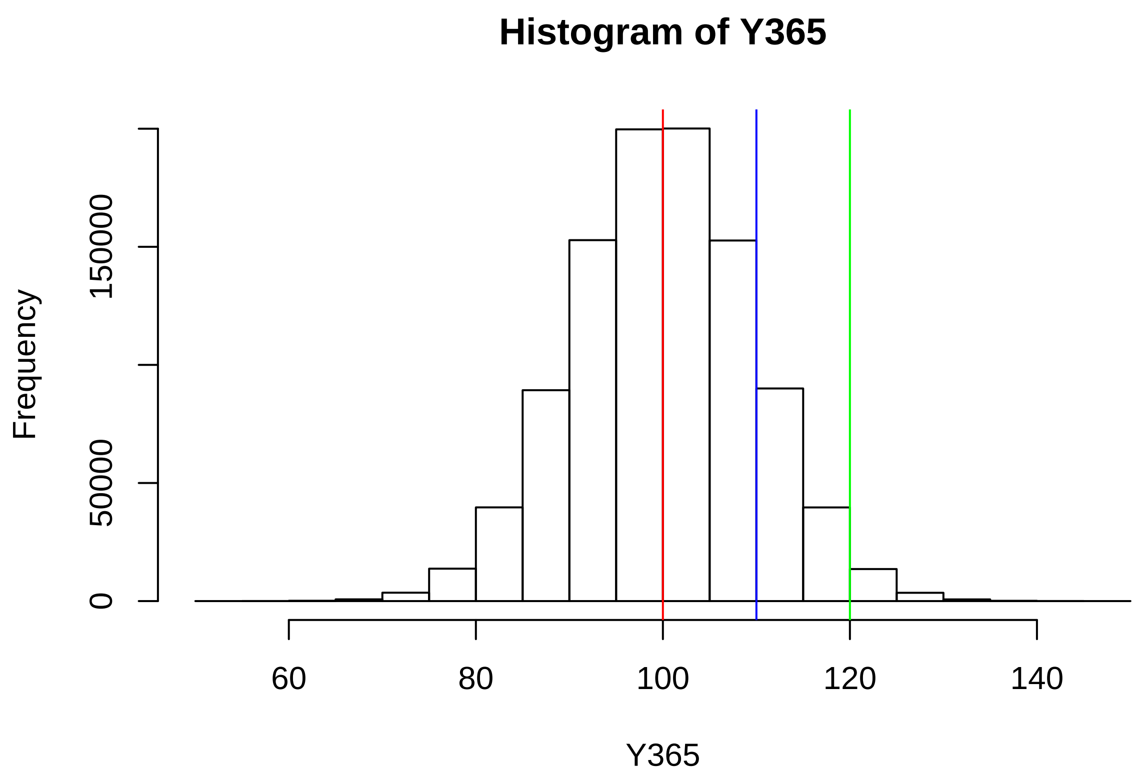
```
##      cut  count  simprob    theory         diff
## [1,] 100 500328 0.500328 0.5000000  0.000328000
## [2,] 110 147564 0.147564 0.1472537  0.000310303
## [3,] 120  17898 0.017898 0.0180158 -0.000117843
```

Across 1,000,000 simulations, the sample mean is $E[Y_{365}] = 100.004022$, which is close to the theoretical value $Y_1 = 100$.

The sample standard deviation is $STDEV[Y_{365}] = 9.536152$, which is close to the theoretical value $\sqrt{91} = 9.539392$.

Histogram

```r
hist(Y365,breaks=21)
abline(v=100,col="red")
abline(v=110,col="blue")
abline(v=120,col="green")
```



---

2. Calculate the expected value and variance of the *binomial* distribution using the moment generating function.

The moment-generating function is defined as $g(t) = E\left[e^{tX}\right] = \sum_{j=1}^{\infty} e^{tx_j} p_X(j)$.

For the binomial, $j \in \{0, 1, 2, ..., n\}$; $x_j = j$; and $p_X(j) = \binom{n}{j} p^j (1-p)^{n-j}$.

Then

$$\begin{aligned} g(t) &= \sum_{j=0}^{n} e^{tj} \binom{n}{j} p^j (1-p)^{n-j} \\ &= \sum_{j=0}^{n} \binom{n}{j} p^j e^{tj} (1-p)^{n-j} \\ &= \sum_{j=0}^{n} \binom{n}{j} \left(pe^t\right)^j (1-p)^{n-j} \\ &= \left(pe^t + (1-p)\right)^n \end{aligned}$$

First derivative:

So, $g' = \frac{dg}{dt} = n\left(pe^t + (1-p)\right)^{n-1} pe^t$

and

$$\begin{aligned} \mu_1 &= g'(t=0) \\ &= n\left(pe^0 + (1-p)\right)^{n-1} pe^0 \\ &= n\,(1)^{n-1}\,p \\ &= np \end{aligned}$$

Expected Value = $\mu_1 = np$

Second derivative:

$g'' = \frac{d^2g}{dt^2} = n(n-1)\left(pe^t + (1-p)\right)^{n-2} (pe^t)^2 + n\left(pe^t + (1-p)\right)^{n-1} pe^t$

$$\begin{aligned} \mu_2 = g''_{(t=0)} &= n(n-1)\left(pe^0 + (1-p)\right)^{n-2} (pe^0)^2 + n\left(pe^0 + (1-p)\right)^{n-1} pe^0 \\ &= n(n-1)\left(p + (1-p)\right)^{n-2} p^2 + n\left(p + (1-p)\right)^{n-1} p \\ &= n(n-1)p^2 + np \\ &= n^2p^2 - np^2 + np \end{aligned}$$

So,

$$\begin{aligned} \sigma^2 &= \mu_2 - \mu_1^2 \\ &= n^2p^2 - np^2 + np - n^2p^2 \\ &= np - np^2 \\ &= np(1-p) \end{aligned}$$

Variance = $\sigma^2 = \mu_2 - \mu_1^2 = np(1-p)$

---

3. Calculate the expected value and variance of the *exponential* distribution using the moment generating function.

Here the moment-generating function is defined as $g(t) = E\left[e^{tX}\right] = \int\limits_{x=0}^{\infty} e^{tx} f_X(x)dx$.

For the exponential, $X \in [0, \infty)$ and $f_X(x) = \lambda e^{-\lambda x}$.

Then

$$
\begin{aligned}
g(t) &= \int\limits_{x=0}^{\infty} e^{tx}\lambda e^{-\lambda x} dx \\
&= \int\limits_{x=0}^{\infty} \lambda e^{(t-\lambda)x} dx \\
&= \left. \frac{\lambda e^{(t-\lambda)x}}{(t-\lambda)} \right|_{x=0}^{x=\infty} \\
&= \frac{0-\lambda}{(t-\lambda)}, \quad for \quad t < \lambda \\
&= \frac{\lambda}{\lambda - t}, \quad for \quad t < \lambda \\
&= \lambda(\lambda - t)^{-1}, \quad for \quad t < \lambda
\end{aligned}
$$

First derivative:

$$
\begin{aligned}
g'(t) &= \frac{dg}{dt} \\
&= -\lambda(\lambda - t)^{-2}(-1) \\
&= \lambda(\lambda - t)^{-2} \\
&= \frac{\lambda}{(\lambda - t)^2}
\end{aligned}
$$

$$
\begin{aligned}
\mu_1 &= g'(0) \\
&= \left. \frac{\lambda}{(\lambda - t)^2} \right|_{t=0} \\
&= \frac{\lambda}{\lambda^2} \\
&= \frac{1}{\lambda} \\
&= \lambda^{-1}
\end{aligned}
$$

Expected value: $\mu_1 = \frac{1}{\lambda}$

Second derivative:

$$
\begin{aligned}
g''(t) &= \frac{d^2 g}{dt^2} \\
&= -2\lambda(\lambda - t)^{-3}(-1) \\
&= 2\lambda(\lambda - t)^{-3} \\
&= \frac{2\lambda}{(\lambda - t)^3}
\end{aligned}
$$

$$
\begin{aligned}
\mu_2 &= g''(0) \\
&= \left. \frac{2\lambda}{(\lambda - t)^3} \right|_{t=0} \\
&= \frac{2\lambda}{\lambda^3} \\
&= \frac{2}{\lambda^2}
\end{aligned}
$$

So,

$$
\begin{aligned}
\sigma^2 &= \mu_2 - \mu_1^2 \\
&= \frac{2}{\lambda^2} - \frac{1}{\lambda^2} \\
&= \frac{1}{\lambda^2} \\
&= \lambda^{-2}
\end{aligned}
$$

Variance = $\sigma^2 = \mu_2 - \mu_1^2 = \frac{1}{\lambda^2}$