Experiment-4

siju.swamy@saintgits.org

14/07/2021

Study of binomial distribution Plots of density and distribution functions. Normal approximation to the Binomial distribution. Central limit theorem.

Normal Approximation of Binomial Distribution Using Central Limit Theorem

Let $\{X_1, \ldots, X_n\}$ be a random sample of size n — that is, a sequence of independent and identically distributed (i.i.d.) random variables drawn from a distribution of expected value given by μ and finite variance given by σ^2. Suppose we are interested in the sample average

$\bar{X}_n \equiv \dfrac{X_1 + \cdots + X_n}{n}$ of these random variables. By the law of large numbers, the sample averages converge almost surely (and therefore also converge in probability) to the expected value μ as $n \to \infty$. The classical central limit theorem describes the size and the distributional form of the stochastic fluctuations around the deterministic number μ during this convergence. More precisely, it states that as n gets larger, the distribution of the difference between the sample average $\bar{X}_n$ and its limit μ, when multiplied by the factor $\sqrt{n}$ (that is $\sqrt{n}(\bar{X}_n - \mu)$) approximates the normal distribution with mean 0 and variance σ^2.

The Central Limit Theorem illustrates that a sample will follow normal distribution if the sample size is large enough. We will use that as well as the rules around determining probabilities in a normal distribution, to arrive at the probability in a Binomial or Poisson case with large n.

Problem 1: I have a group of 100 friends who are smokers. The probability of a random smoker having lung disease is 0.3. What are chances that a maximum of 35 people wind up with lung disease?

Solution: Here we have to find $P(x \leq 35)$. Using a Binomial distribution we can calculate it as:

```
pbinom(35,size=100,prob=0.3)
```

```
## [1] 0.8839214
```

To use a normal distribution, first we have to find $\mu = np = 100 \times 0.3 = 30$, $\sigma = \sqrt{npq} = \sqrt{100 \times 0.3 \times 0.7} = 4.582576$.

```
pnorm(35,30,4.582576)
```

```
## [1] 0.8623832
```

Problem 2: Suppose 35% of all households in Carville have three cars, what is the probability that a random sample of 80 households in Carville will contain at least 30 households that have three cars.

Solution: Here we have to find $p(x \geq 30)$.

Using Binomial distribution,

```
1-pbinom(30,80,0.35)
```

```
## [1] 0.2764722
```

For using Normal distribution $\mu = 28, \sigma = 4.266146$.

```r
pnorm(30,28,4.27,lower.tail = F)
```

```
## [1] 0.319755
```

Problem 3: About two out of every three gas purchases at Cheap Gas station are paid for by credit cards. 480 customers buying gas at this station are randomly selected. Find the following probabilities using the binomial distribution, normal approximation and using the continuity correction.

1. Find n, p, q, the mean and the standard deviation.
2. Find the probability that greater than 300 will pay for their purchases using credit card.
3. Find the probability that between 220 to 320 will pay for their purchases using credit card.
4. Generate a random number using the command floor(rnorm(1, mean=200, sd=50)).

Write this number down. Lets call it N. (This number will be different for each student.)

5. Find the probability that at most N customers will pay for their purchases using credit card.