Stats 314, Data Analysis #1

Cody Malick
malickc@oregonstate.edu

October 9, 2016

Part I

a

Uniform distribution, because it is continuous, each time has an equal probability, and distribution. This is a property unique to uniform distributions.

b

Poisson distribution, because we are given an average rate over a discrete random variable.

c

Exponentional distribution, we're given a continuous random variable with an average rate (8), while also being given a skew indicating an exponential distribution.

d

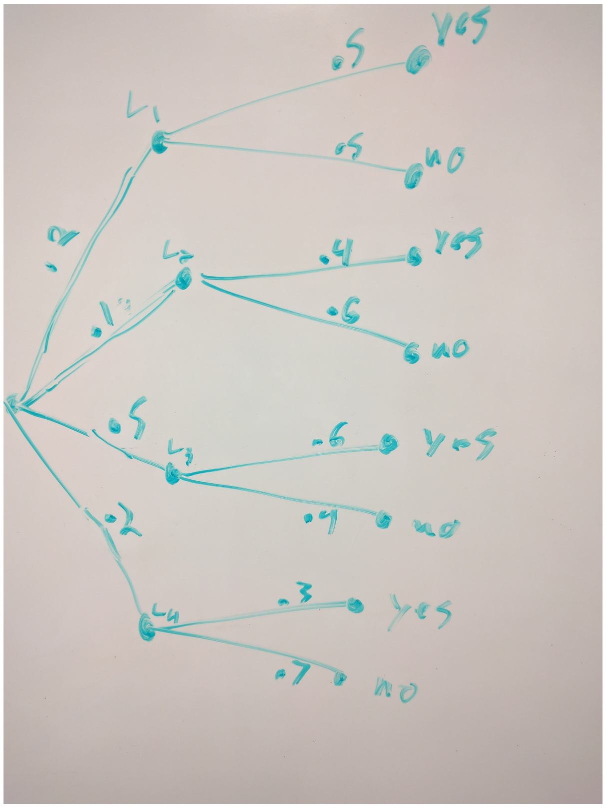
Binomial distribution, as we are given a set of independent events, two possible outcomes, a probability of failure (.25) and through that success (.75).

e

Normal distribution, as we are given an average value (6), the standard deviation from that value (.6), and are told that values farther from the median value are less likely. This describes a typical bell curve.

Part II

a



b

To get the probability of a ticket, we have to use the law of total probability to calculate the combined probability of getting a ticket in each area given that they are in operation.
L1: $.2 * .5 = .1$
L2: $.1 * .4 = .04$
L3: $.5 * .6 = .3$
L4: $.2 * .3 = .06$

P(ticket)$= .1 + .04 + .3 + .06 = .5$

Part III

a

HHHH
HHHT
HHTH
HTHH
THHH
HHTT
HTHT
THHT
THTH
HTTH
TTHH
HTTT

THTT
TTHT
TTTH
TTTT

b

$4/16 = .25$

c

$5/16 = .3125$

d

Binomial distribution, because we have a set of independent events with a discrete random variable, and we're interested in the success rate for specific outcomes.

e

To get the probability of $X = 3$, we have to take:
$P(X = 3) = 4/16 = .25$

f

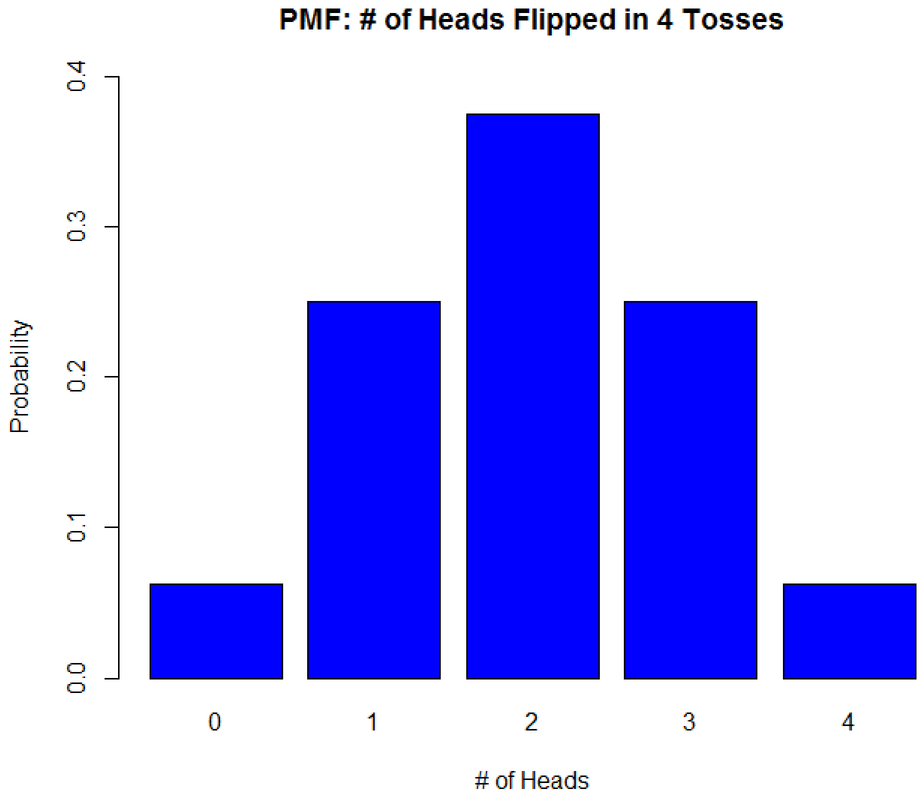
To get the probability of $X >= 3$, we have to take:
$1 - P(X >= 3) = 1 - (P(0) + P(1) + P(2)) = 1 - .6875 = .3125$

g

They sure are!

h

The most likely number is 2. The likelihood is .375



i

$E(X) = np$, where n is the number of coin tosses, and p is the probability of success.
$E(X) = 4 * .5 = 2$, which is our most likely number of heads/tails.

Part IV

a

To verify the distribution $f(x)$, we can simply integrate the two equations from negative infinity to infinity get the range of the values we're interested in.

For anything other than $0 < x < 1$, we can simply plug in zero, so that simplifies the problem. We can take the integral of $5(1-x)^4$, from zero to one because we know everything less than zero and greater than one equal zero.

$\int_{-\infty}^{\infty} 5(1-x)^4 dx$

$= \int_0^1 5(1-x)^4 dx$

Using u-substitution:
$5\int_0^1 u^4 du, u = (x-1), du = dx$

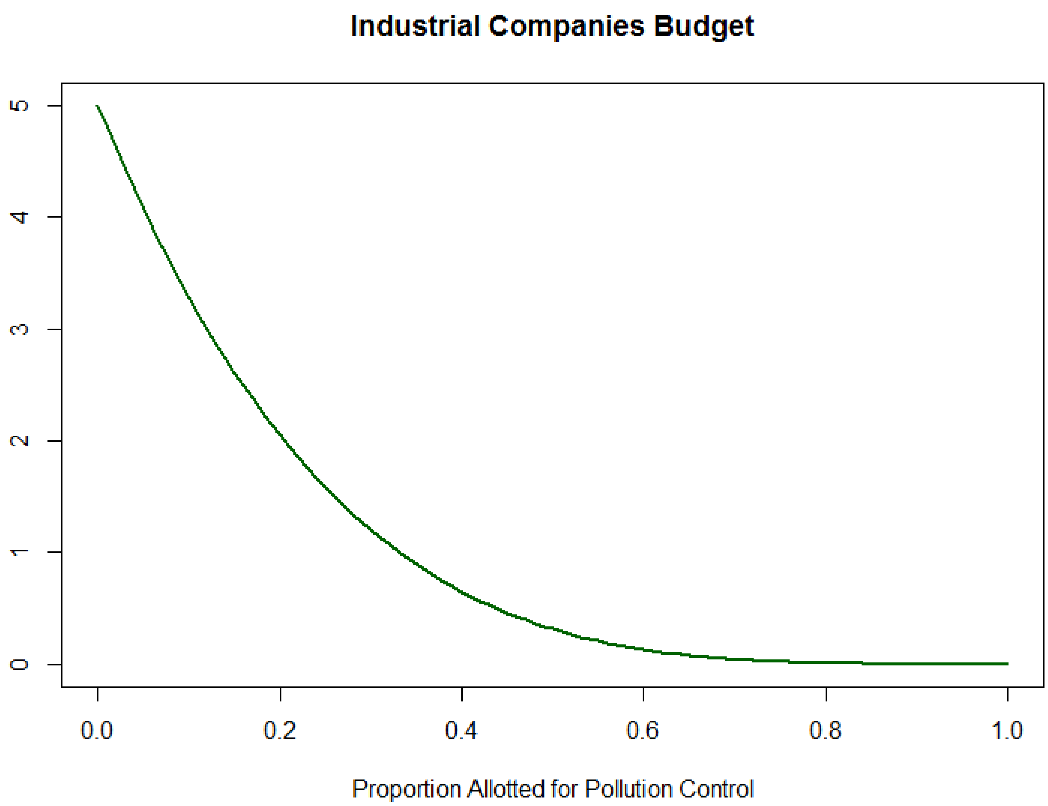
$$= \frac{5(x-1)^5}{5} = (x-1)^5\Big|_0^1$$

$= (1-1)^5 - (0-1)^5 = 0 - (-1^5) = 1$

This answer justifies our probability mass function because the total probability equals 1.

b

The shape of the distribution is an inverse exponential curve. It is not, however, an exponential distribution. There are no e values in the equation. Starting at five, decreasing towards zero.



c

The expected value of this distribution is $E(X) = \int_{-\infty}^{\infty} x * f(x)dx$. Again, we can simply take the integral from 0 to 1 because of the bounds of our integration.

$E(X) = \int_{-\infty}^{\infty} x * 5(x-1)^4 dx$

$= \int_0^1 x5(x-1)^4 dx = 5\int_0^1 x(x-1)^4 dx$

$= 5\int_0^1 (x-1)^4 * (x-1+1) = 5\int_0^1 (x-1)^5 + 5\int_0^1 (x-1)^4$

Using u-substitution, where $u = x - 1$:
$= 5\int_0^1 u^5 du + = 5\int_0^1 u^4 du = \frac{5(x-1)^6}{6} + (x-1)^5\Big|_0^1$

$= (\frac{5(1-1)^6}{6} + (1-1)^5) - (\frac{5(0-1)^6}{6} + (0-1)^5 = 0 - (\frac{5}{6} - 1)$

$= -(-\frac{1}{6}) = \frac{1}{6}$

So the expected value of our function is $\frac{1}{6}$

d

The standard deviation of the function is: $SD = \sqrt{V(X)}$. So we need to find $V(X)$.
$V(X) = E(X^2) - E(X) = E(X^2) - \frac{1}{6}$

$E(X^2) = \int_0^1 (x^2)5(x-1)^4 dx = 5\int_0^1 (x^2)(x-1)^4 dx$

Using u-substitution, where $u = x - 1$:
$= 5\int_0^1 u^4(u+1)^2 = 5\int_0^1 u^6 + 2u^5 + u^4 du$

$= 5\int_0^1 u^6 du + 5\int_0^1 2u^5 du + 5\int_0^1 u^4 du$

$= 5(\frac{(x-1)^7}{7} + \frac{2(x-1)^6}{6} + \frac{(x-1)^5}{5})\Big|_0^1$

$= \frac{5(x-1)^7}{7} + \frac{10(x-1)^6}{6} + (x-1)^5\Big|_0^1$

$= (\frac{5(1-1)^7}{7} + \frac{10(1-1)^6}{6} + (1-1)^5) - (\frac{5(0-1)^7}{7} + \frac{10(0-1)^6}{6} + (0-1)^5)$

$= (0) - (\frac{-5}{7} + \frac{10}{6} - 1 = -(-.71428 + 1.6666 - 1) = .0476$

With our $E(X^2)$ we can calculate the variance and the standard deviation:

$V(X) = .0476 - .02778 = .0198$

$SD(X) = \sqrt{(}.0198) = .1408$

e

The CDF is the integral of the PDF from 0 to x, which we can quickly calculate given previous work:

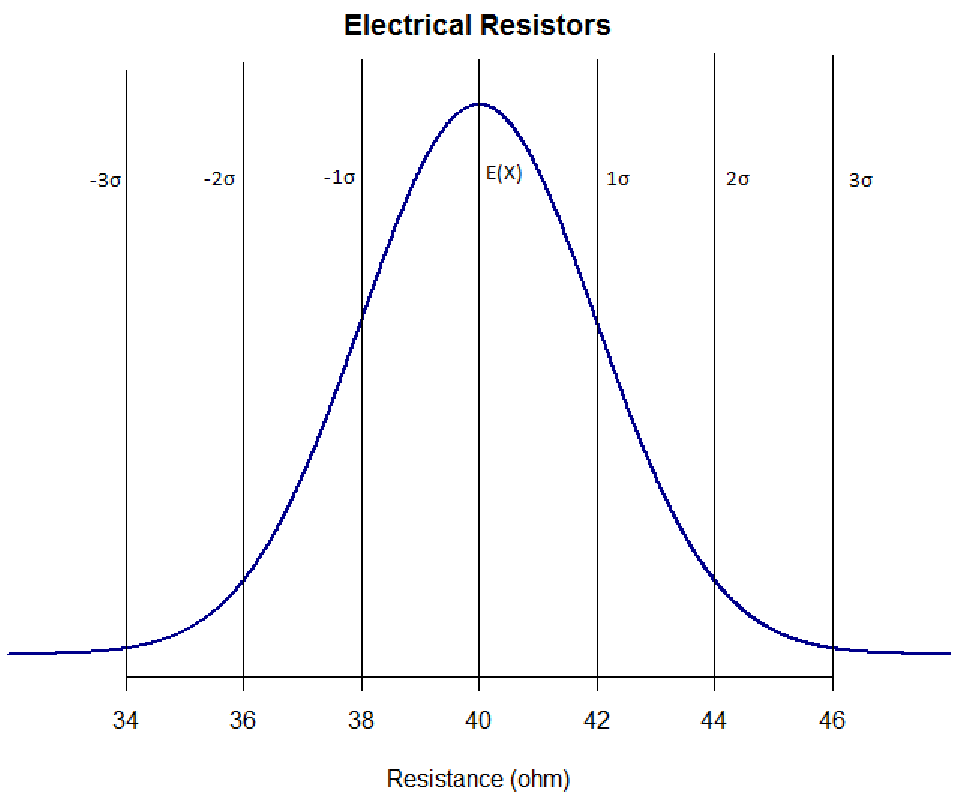
$\int_0^x f(x)dx = 5\int_0^x (1-x)^4 = (x-1)^5\Big|_0^x = (x-1)^5 - (0-1)^5 = (x-1)^5 + 1$
Our CDF then is: $F(x) = (x-1)^5 + 1$

f

We can find this quickly by plugging in .05 into our equation:
$(.05 - 1)^5 + 1 = .2262$

Part V

a



b

99.7%

c

At 43, we are 1.5 standard deviations from the median. With that info, we can jump straight to the z-table. With a z-score of -1.5, or 1-zscore(1.5), then the answer is .06681

d

To find the percentile, we have to find what zscore occurs at about .25, which in this case means $-.67 * \sigma + \mu = -.67 * 2 + 40 = 38.66$