Notebook

March 22, 2019

Local date & time is : 03/22/2019 16:08:33 PDT

1 newpage

1.1 Part 1: The Operations

We will start with simple examples of the basic operations and syntax. For more, take a look at the SymPy Tutorial.

2 newpage

2.1 Part 2: Working with Densities

Now that you have an idea of what you can do with `SymPy`, it's time for some probability theory.

In Prob 140, calculus is used almost exclusively when working with densities. Later in the course it will be used for some optimization as well, but density calculations are its main use.

In this part of the lab, you will use `SymPy` to work with a very simple density. That way, you will be able to check by hand that `SymPy` is getting the right answers. In subsequent parts of the lab, the densities will get more complicated. Treat this part of the lab as a warm-up.

```
In [32]: x = Symbol('x', positive=True)
         f_X = 6*x**5
         f_X
```

Out[32]:

$$6x^5$$

the integral from 0 to 1 of our fx should be 1.

```
In [33]: Integral(f_X, (x, 0 , 1)).doit()
```

Out[33]:

$$1$$

```
In [35]: # Display the integral for P(X > 0.75)

         Integral(f_X, (x, 0.75, 1))
```

Out[35]:

$$\int_{0.75}^{1} 6x^5\, dx$$

```
In [36]: # Evaluate P(X > 0.75)

         Integral(f_X, (x, 0.75, 1)).doit()
```

Out[36]:

$$0.822021484375$$

Your answer
First blank: 0.45
Second blank: 0.95

```
In [37]: # P(|X - 0.7| < 0.25)

         Integral(f_X, (x, 0.45, 0.95)).doit()
```

Out[37]:

$$0.726788125$$

```
In [38]: # E(X)

         expectation_X = Integral(x*f_X, (x, 0, 1)).doit()
         expectation_X
```

Out[38]:

$$\frac{6}{7}$$

In [39]:
```
# SD(X)

variance = Integral(x**2*f_X, (x, 0, 1)).doit() - (Integral(x*f_X, (x, 0, 1)).doit())**2
SD = variance**0.5
SD
```

Out[39]:

$$0.123717914826348$$

In [40]:
```
# Display the integral for E(log(X))

Integral(log(x)*f_X, (x, 0, 1))
```

Out[40]:

$$\int_0^1 6x^5 \log(x)\, dx$$

In [41]:
```
# Numerical value of E(log(X))

Integral(log(x)*f_X, (x, 0, 1)).doit()
```

Out[41]:

$$-\frac{1}{6}$$

Log value is negative from 0 to 1. So he value of log(x)*6*x**5 is also negative for x in 0, 1. so summing up the value of log(x)*6*x**5 and dx, the sum is negative

In [42]:
```
# Display the integral for E(sin_squared(X))

Integral(sin(x)**2*f_X, (x, 0, 1))
```

Out[42]:

$$\int_0^1 6x^5 \sin^2(x)\, dx$$

In [43]:
```
value_of_integral = Integral(sin(x)**2*f_X, (x, 0, 1)).doit()
value_of_integral
```

Out[43]:

$$-\frac{21}{2}\sin(1)\cos(1) + 8\cos^2(1) + \frac{17}{4}\sin^2(1)$$

In [45]:
```
g = pi*x**2
g
```

Out[45]:

$$\pi x^2$$

It is from 0 to pi.

In [47]:
```
v = Symbol('v', positive=True)
g_inverse = solve(g - v, x)
g_inverse
```

Out[47]:

$$\left[\frac{\sqrt{v}}{\sqrt{\pi}} \right]$$

In [49]:
```
deriv_g = diff(g, x)
deriv_g
```

Out[49]:

$$2\pi x$$

In [50]:
```
"""For v in the interval 0 to pi, the density of V at the point v is:"""

f_V = (f_X / deriv_g).subs(x, g_inverse)
f_V
```

Out[50]:

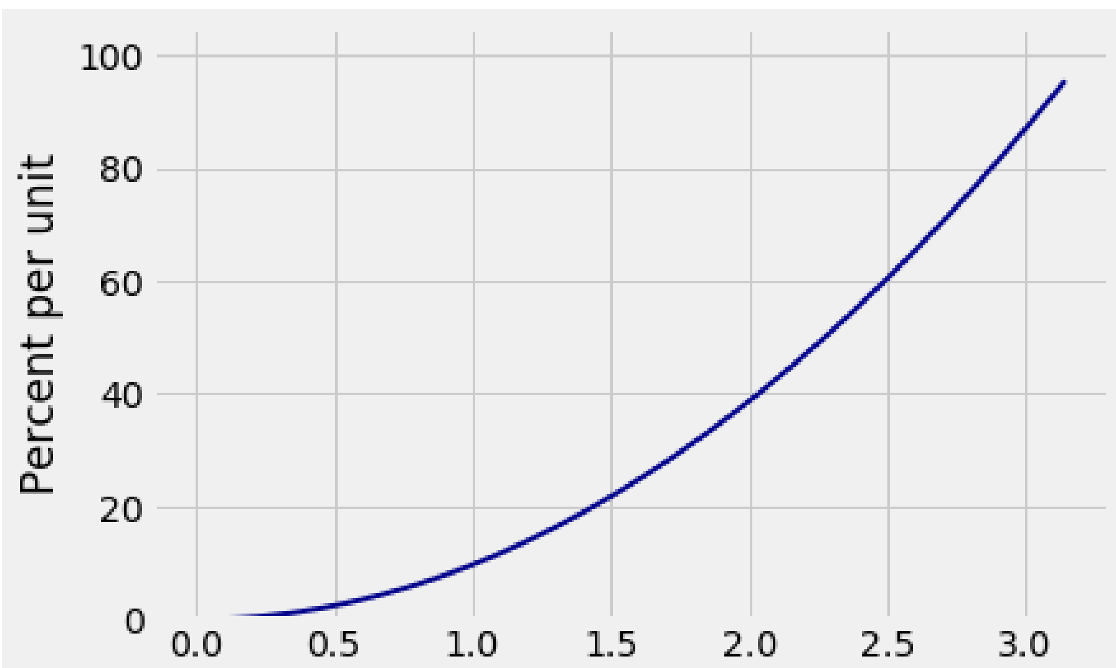
$$\frac{3v^2}{\pi^3}$$

In [51]:
```
Integral(f_V, (v, 0, pi)).doit()
```

Out[51]:

$$1$$

In [52]:
```
Plot_continuous([0, np.pi], f_V)
```



3 newpage

3.1 Part 3: Transforming an$F$Distribution

The density you worked with in Part 2 was deliberately chosen to be simple. The real value of SymPy is that without much difficulty you can use essentially the same code to handle more complicated density functions.

In this part of the lab you will start with one of the famous distributions of statistical inference, confusingly named the$F$distribution. It's not named for our old friend the cdf. It is named for our other old friend Sir Ronald Fisher, the extraordinary scientist whose contributions include what is now the standard method of testing statistical hypotheses. You might recall from Data 8 that it was Fisher who found 5% to be a "convenient" cutoff for the P-value, which then became set in stone as the cutoff for statistical significance.

The$F$distribution arises as the distribution of the ratio of two independent gamma random variables (apart from a constant multiplier). The ratio arises in tests of whether three or more random samples come from the same underlying distribution.

For the purposes of this lab, the$F$distribution is just an ordinary distribution on the positive numbers. It has two parameters, which we will call$n$for "numerator" and$d$for "denominator". And its density has a rather intimidating formula.

```
In [122]: def constant_F(n, d):
              return gamma(n/2 + d/2)/(gamma(n/2)*gamma(d/2))*(n/d)**(n/2)
```

```
In [125]: x = Symbol('x', positive=True)
          f_X = constant_F(6,4)*x**2*(1+6/4*x)**-5
          f_X
```

Out[125]:

$$\frac{40.5x^2}{(1.5x+1)^5}$$

```
In [126]: Integral(f_X, (x, 0, oo)).doit()
```

Out[126]:

$$1.0$$

The possible values of V is 0 to 1, not including 0 and 1

```
In [128]: g = 6/4*x / (1+6/4*x)
          g
```

Out[128]:

$$\frac{1.5x}{1.5x+1}$$

g is increasing. When x is 0 g(0) = 0, and when x -> inf, g(x) converges to 1.

```
In [129]: v = Symbol('v', positive=True)
          g_inverse = solve(g - v, x)[0]
          g_inverse
```

Out[129]:

$$-\frac{2.0v}{3.0v-3.0}$$

for possible value of v in [0, 1) 3.0v - 3.0 always have a negtive value. Hence the overall evaluation of g inverse is positive. we can disregard the case when v = 1, because as x goes to infinity, v gets really close to 1, but is never actually equal to 1.

```
In [130]: deriv_g = diff(g, x)
          deriv_g
```

Out[130]:

$$-\frac{2.25x}{(1.5x+1)^2}+\frac{1.5}{1.5x+1}$$

simplifying the derivative, there is no negative signs or subtraction. Only addition is included. Hence for any value of positive x, derivative is positive.

```
In [131]: # Density of V:

          """For v in the interval 0 to 1, the density of V at the point v is:"""

          f_V = (f_X / deriv_g).subs(x, g_inverse)
          f_V
```

Out[131]:

$$\frac{162.0v^2}{(3.0v-3.0)^2\left(-\frac{3.0v}{3.0v-3.0}+1\right)^5\left(\frac{4.5v}{(3.0v-3.0)\left(-\frac{3.0v}{3.0v-3.0}+1\right)^2}+\frac{1.5}{-\frac{3.0v}{3.0v-3.0}+1}\right)}$$

```
In [132]: f_V = factor(simplify(f_V))
          f_V
```

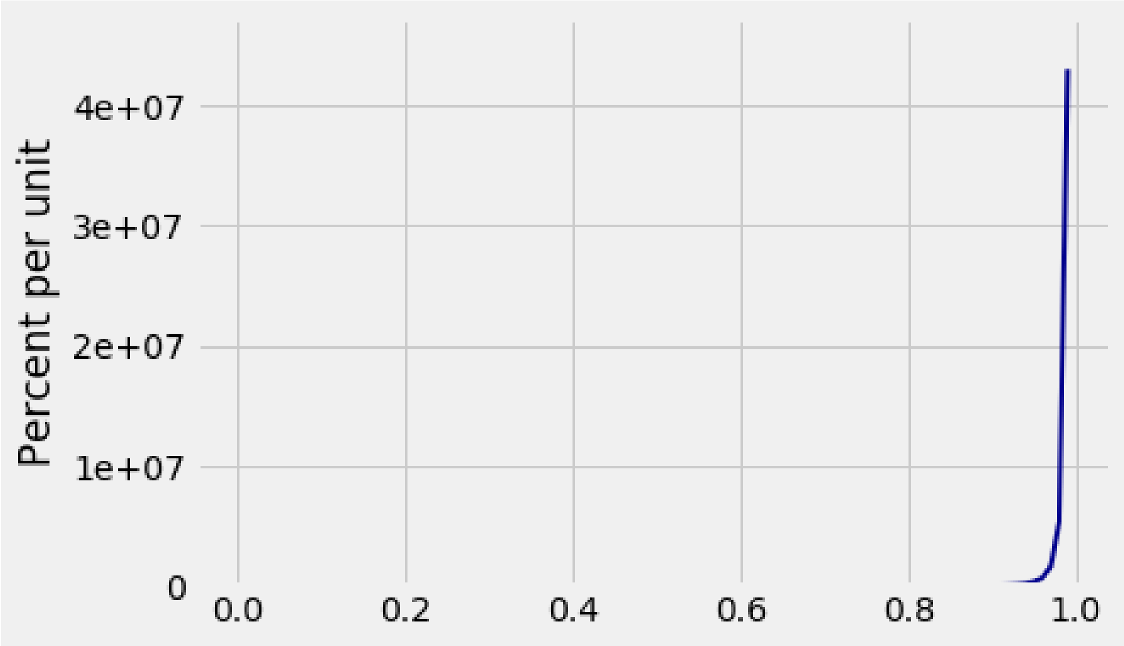
Out[132]:

$$-12.0v^2(v-1)$$

This is a beta density, with r =

```
In [111]: Plot_continuous([0,1], f_V)
```

```
/srv/app/venv/lib/python3.6/site-packages/numpy/__init__.py:1: RuntimeWarning: divide by zero encountere
  """
```



4 newpage

4.1 Part 4: Working with Joint Densities

A joint density is a function of two variables. This part of the lab shows you how to work with such functions in SymPy.

Before you get started, please read Section 17.1 of the textbook. The rest of the lab will go faster that way.

Let$X$and$Y$have joint density given by

$$f(x,y) \;=\; x+y, \quad 0 < x,y < 1$$

In order to find probabilities, densities, or expectations of functions of$X$and$Y$, you have to create two symbols, which we will assign to the names x and y. Since both have values in$(0,1)$, we will use the positive=True option.

In [68]:
```
# P(Y > 2X) by integrating y first and then x

Integral(f, (y, 2*x, 1), (x, 0, 1/2))
```

Out[68]:

$$\int_{0}^{0.5}\int_{2x}^{1} (x+y)\; dy\, dx$$

In [69]:
```
# P(Y > 2X) by integrating x first and then y

Integral(f, (x, 0, y/2 ), (y, 0, 1))
```

Out[69]:

$$\int_{0}^{1}\int_{0}^{\frac{y}{2}} (x+y)\; dx\, dy$$

In [70]:
```
# numerical value of P(Y > 2X)

Integral(f, (y, 2*x, 1), (x, 0, 1/2)).doit(),Integral(f, (x, 0, y/2 ), (y, 0, 1)).doit()
```

Out[70]:

$$\left(0.208333333333333, \quad \frac{5}{24}\right)$$

In [71]:
```
5/24
```

Out[71]:

0.20833333333333334

In [77]:
```
# For x in the interval 0 to 1, the marginal density of X at the point x is:

f_X = Integral(f, (y, 0, 1)).doit()
f_X
```

Out[77]:

$$x+\frac{1}{2}$$

The marginal density of Y is same as marginal density of X, only x substituted for y. x and y have same inteval, and the density function is symmetric for x and y. Hence the integral in terms of x and y should be the same.

E(X+Y) should be 2 times E(X)

In [78]:
```
Integral(2*x*f_X, (x, 0, 1 )).doit()
```

Out[78]:

$$\frac{7}{6}$$

In [74]:
```
# E(XY) displayed as an integral

Integral(x*y*f, (y, 0 ,1), (x, 0 , 1))
```

Out[74]:

$$\int_0^1 \int_0^1 xy\,(x+y)\;dy\,dx$$

In [76]:
```
# Numerical value of E(XY)

Integral(x*y*f, (y, 0 ,1), (x, 0 , 1)).doit()
```

Out[76]:

$$\frac{1}{3}$$

In [79]:
```
"""Given X=x, the possible values of Y are 0 to 1 For y in this range:"""

f_Y_given_X_is_x = f / (f_X)
f_Y_given_X_is_x
```

Out[79]:

$$\frac{x+y}{x+\frac{1}{2}}$$

In [80]:
```
Integral(f_Y_given_X_is_x, (y, 0 ,1))
```

Out[80]:

$$\int_0^1 \frac{x+y}{x+\frac{1}{2}}\,dy$$

In [82]:
```
simplify(Integral(f_Y_given_X_is_x, (y, 0 ,1)).doit())
```

Out[82]:

$$1$$

In [84]:
```
f_Y_given_X_is_025 = f_Y_given_X_is_x.subs(x, 0.25)
f_Y_given_X_is_025
```

Out[84]:

$$1.33333333333333y + 0.333333333333333$$

In [86]:
```
prob = Integral(f_Y_given_X_is_025, (y, 0.3, 1)).doit()
cond_exp = Integral(y*f_Y_given_X_is_025, (y, 0, 1)).doit()

prob, cond_exp
```

Out[86]:

$$(0.84, \quad 0.611111111111111)$$