

## BST 140.651 Problem Set 2

Kai Aragaki

9/15/2020

### Problem 1

---

$$\text{Var}(X) = E[X^2] - E[X]^2$$

Where  $\text{Var}(X) = E[(X - \mu)^2]$ , so

$$E[(X - \mu)^2] = E[X^2] - E[X]^2$$

$$E[X^2 - 2X\mu + \mu^2] = E[X^2] - E[X]^2$$

$$E[X^2] + E[-2X\mu] + E[\mu^2] = E[X^2] - E[X]^2$$

Since  $\mu$  is a constant, and constants can be pulled out of the expected value function,

$$E[X^2] - 2\mu E[X] + \mu^2 E[1] = E[X^2] - E[X]^2$$

Simplifying, and with the knowledge that  $E[X] = \mu$ :

$$-2\mu^2 + \mu^2 = -\mu^2$$

From which we find

$$-2\mu^2 = -2\mu^2$$

Which we know to be true.

### Problem 2

---

$a$

---

$g$  is a valid density if all values are positive and if the integral across all values is 1.

We know all values are positive because it is the sum of three densities (which themselves must always be positive), each of which is multiplied by a scalar that is also positive.

To show that the integral across all values is equal to 1,

$$\int_{-\infty}^{\infty} g(x) = 1$$

Since we know

$$g(x) = \pi_1 f_1(x) + \pi_2 f_2(x) + \pi_3 f_3(x)$$

then

$$\int_{-\infty}^{\infty} g(x) = \int_{-\infty}^{\infty} (\pi_1 f_1(x) + \pi_2 f_2(x) + \pi_3 f_3(x))$$

which is equivalent to

$$\int_{-\infty}^{\infty} g(x) = \pi_1 \int_{-\infty}^{\infty} f_1(x) + \pi_2 \int_{-\infty}^{\infty} f_2(x) + \pi_3 \int_{-\infty}^{\infty} f_3(x)$$

and since each function  $f_i$  is a valid density their integrals must equal 1, giving

$$\int_{-\infty}^{\infty} g(x) = \pi_1 + \pi_2 + \pi_3$$

Since we are additionally told  $\sum_{i=1}^3 \pi_i = 1$

$$\int_{-\infty}^{\infty} g(x) = 1$$

■

*b*

We know that mean (or expected value) for a continuous distribution is given by

$$\int_{-\infty}^{\infty} xg(x)$$

Which is

$$\int_{-\infty}^{\infty} x(\pi_1 f_1(x) + \pi_2 f_2(x) + \pi_3 f_3(x))$$

which can be expanded to

$$\pi_1 \int_{-\infty}^{\infty} x f_1(x) + \pi_2 \int_{-\infty}^{\infty} x f_2(x) + \pi_3 \int_{-\infty}^{\infty} x f_3(x)$$

And since we know the definition of the expected value, we can rewrite this as

$$E[X] = \pi_1 \mu_1 + \pi_2 \mu_2 + \pi_3 \mu_3$$

*c*

We have previously shown that the definition of

$$Var(x) = E[X^2] - E[X]^2$$

We can represent the variance of  $g$  in these terms:

$$\int_{-\infty}^{\infty} x^2 g(x) - \left[ \int_{-\infty}^{\infty} x g(x) \right]^2$$

The right term we have previously found - it's simply the square of the expectation value we found in the previous problem. The left terms can be found by first writing  $g$  in terms of  $f$

$$\int_{-\infty}^{\infty} x^2 g(x) = \int_{-\infty}^{\infty} x^2 (\pi_1 f_1(x) + \pi_2 f_2(x) + \pi_3 f_3(x))$$

or distributed:

$$\int_{-\infty}^{\infty} x^2 g(x) = \pi_1 \int_{-\infty}^{\infty} x^2 f_1(x) + \pi_2 \int_{-\infty}^{\infty} x^2 f_2(x) + \pi_3 \int_{-\infty}^{\infty} x^2 f_3(x)$$

and then noticing that the variance equation can be written as follows:

$$\sigma^2 = E[X^2] - \mu^2$$

And solving for  $E[X^2]$ :

$$E[X^2] = \mu^2 + \sigma^2$$

which is equivalent to

$$\int_{-\infty}^{\infty} x^2 f_i(x) = \mu_i^2 + \sigma_i^2$$

Substituting in, we get

$$\int_{-\infty}^{\infty} x^2 g(x) = \pi_1(\mu_1^2 + \sigma_1^2) + \pi_2(\mu_2^2 + \sigma_2^2) + \pi_3(\mu_3^2 + \sigma_3^2)$$

Combining this left hand term with the right hand term, we get

$$Var(x) = \pi_1(\mu_1^2 + \sigma_1^2) + \pi_2(\mu_2^2 + \sigma_2^2) + \pi_3(\mu_3^2 + \sigma_3^2) - (\pi_1\mu_1 + \pi_2\mu_2 + \pi_3\mu_3)^2$$

### Problem 3

a

The mean of a continuous distribution is

$$\int_{lower}^{upper} xf(x)$$

In this case the lower and upper bounds are given as 0 and 1 respectively.

Therefore the mean is given by

$$\begin{aligned} & \int_0^1 (k+1)x^{k+1} \\ & \int_0^1 kx^{k+1} + \int_0^1 x^{k+1} \\ & \left. \frac{kx^{k+2}}{k+2} \right|_0^1 + \left. \frac{x^{k+2}}{k+2} \right|_0^1 \\ & \frac{k}{k+2} + \frac{1}{k+2} \\ & \frac{k+1}{k+2} \end{aligned}$$

b

The variance of this distribution is

$$Var(x) = E[X^2] - E[X]^2$$

We know  $E[X] = \mu$ , which is what we found in part a, so squaring it is trivial. Additionally,

$$E[X^2] = \int_0^1 x^2((k+1)x^k) = \int_0^1 (k+1)x^{k+2}$$

And by distributing,

$$\int_0^1 kx^{k+2} + \int_0^1 x^{k+2}$$

We can see by analogy to the first problem that this will integrate and simplify to

$$\frac{k+1}{k+3}$$

### Problem 4

---

a

---

To calculate the mean, we know that the mean is, for the limits of this given function

$$E[X] = \int_0^{\infty} xf(x) = \frac{1}{3.3} \int_0^{\infty} xe^{\frac{-x}{3.3}}$$

We can integrate by parts using - that is, we know the following to be true:

$$\int u dv = uv - \int v du$$

so if we set  $u = x$ ,  $du = dx$ ,  $v = -3.3e^{\frac{-x}{3.3}}$ , and  $dv = e^{\frac{-x}{3.3}} dx$ ,

$$\int xe^{\frac{-x}{3.3}} = 3.3xe^{\frac{-x}{3.3}} + 3.3 \int e^{\frac{-x}{3.3}} dx = 3.3xe^{\frac{-x}{3.3}} - 3.3^2 e^{\frac{-x}{3.3}}$$

Substituting this integral back into the original, we obtain

$$xe^{\frac{-x}{3.3}} - 3.3e^{\frac{-x}{3.3}}$$

evaluating at its limits of 0 and  $\infty$ , and with the knowledge that  $e^{-x}$  approaches 0 much faster than  $x$  approaches infinity,

$$xe^{\frac{-x}{3.3}} - 3.3e^{\frac{-x}{3.3}} \Big|_0^{\infty} = [0 + 0] - [0 - 3.3] = 3.3 \quad \blacksquare$$

To calculate the variance, we solve the following:

$$Var(X) = \int_0^{\infty} x^2 f(x) - E[X]^2$$

But since we know the value of  $E[X]$  (and squaring it is trivial) we focus our efforts on the left hand side of the equation:

$$\int_0^{\infty} x^2 f(x) = \frac{1}{3.3} \int_0^{\infty} x^2 e^{\frac{-x}{3.3}}$$

As before, we begin to integrate by parts, with  $u = x^2$ ,  $du = 2xdx$ ,  $v = -3.3e^{\frac{-x}{3.3}}$ , and  $dv = e^{\frac{-x}{3.3}}$ :

$$\frac{1}{3.3} [-3.3x^2 e^{\frac{-x}{3.3}} + 6.6 \int e^{\frac{-x}{3.3}} x dx]$$

We integrate by parts again, this time where  $u = x$ ,  $du = dx$ ,

$v = -3.3e^{\frac{-x}{3.3}}$ , and  $dv = e^{\frac{-x}{3.3}} dx$ ,

$$\frac{1}{3.3} \left[ -3.3x^2 e^{\frac{-x}{3.3}} + 6.6 [-3.3xe^{\frac{-x}{3.3}} + 3.3 \int e^{\frac{-x}{3.3}} dx] \right]$$

Which is equivalent to

$$\frac{1}{3.3} \left[ -3.3x^2 e^{\frac{-x}{3.3}} + 6.6 [-3.3xe^{\frac{-x}{3.3}} - 3.3^2 e^{\frac{-x}{3.3}}] \right]$$

As before, evaluating at the limits 0 to  $\infty$  and substitute back in  $E[X]^2$

$$Var(X) = 2(3.3)^2 - 3.3^2 = 3.3^2 = 10.89 \quad \blacksquare$$

b

---

If we substitute 3.3 for  $\beta$  in our previous solutions, we find

$$E[X] = xe^{\frac{-x}{\beta}} - \beta e^{\frac{-x}{\beta}} \Big|_0^{\infty} = [0 + 0] - [0 - \beta] = \beta$$

and

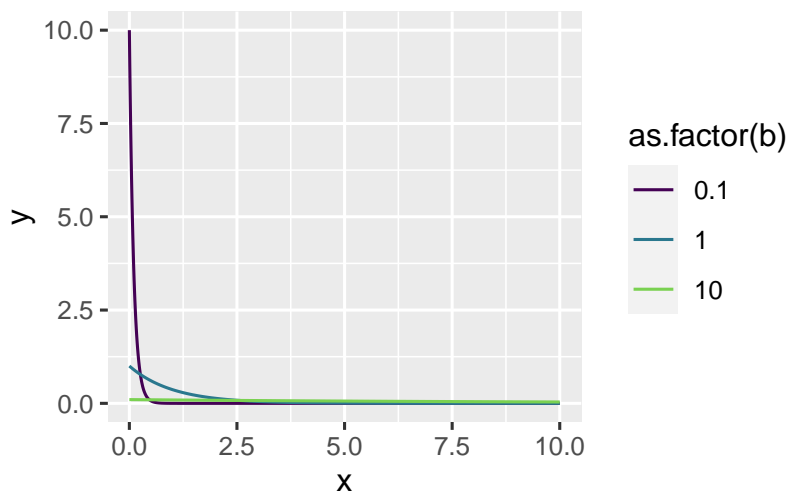
$$\text{Var}(X) = \frac{1}{\beta} \left[ -\beta x^2 e^{\frac{-x}{\beta}} + 2\beta [-\beta x e^{\frac{-x}{\beta}} - \beta^2 e^{\frac{-x}{\beta}}] \right] - \beta^2 = \beta^2$$

*c*

---

```
tib <- expand_grid(x = seq(0, 10, by = 0.01), b = c(0.1, 1, 10)) %>% mutate(y = (1/b) *
  exp(-x/b))

ggplot(tib, aes(x = x, y = y, color = as.factor(b))) + scale_color_viridis_d(end = 0.8) +
  geom_line()
```



*Problem 5*

*a*