Stat 322 - Solutions to Homework 7

1 Apple-Tree Farm

Able and Baker are both apple-tree farmers (they grow apple trees). Assume, that apple trees grow according to a normal distribution.
On the Able Farm, trees grow with a mean of 1m per year and a standard deviation of 25 cm. Baker manages to get an average growth of 1.1m per year with a standard deviation of 35cm.
For the following questions again state each time which random variable you use and what distribution assumption you make.

(a) What is the probability for a tree on the Able Farm to grow more than 1.2m in a year? what if the tree was on the Baker Farm?

Let X be the growth of a tree on the Able farm in one year; then $X \sim N_{\mu=1,\sigma^2=0.25^2}$

$$\begin{aligned} P(X > 1.2) &= 1 - P(X \leq 1.2) = 1 - N_{\mu=1,\sigma^2=0.25^2}(1.2) \overset{standardization}{=} \\ &= 1 - N_{0,1}(\frac{1.2-1}{0.25}) = 1 - N_{0,1}(0.8) \overset{table}{=} 1 - 0.7881 = 0.2119. \end{aligned}$$

Analogously, for Y := growth of a tree on the Baker farm in one year; then $Y \sim N_{\mu=1.1,\sigma^2=0.35^2}$

$$\begin{aligned} P(Y > 1.2) &= 1 - P(Y \leq 1.2) = 1 - N_{\mu=1.1,\sigma^2=0.35^2}(1.2) \overset{standardization}{=} \\ &= 1 - N_{0,1}(\frac{1.2-1.1}{0.35}) \approx 1 - N_{0,1}(0.29) \overset{table}{=} 1 - 0.6141 = 0.3859. \end{aligned}$$

(b) What is the probability for a tree on the Able Farm to grow less than .8m in a year? what if the tree was on the Baker Farm?

Using the same random variables X and Y as above:

$$\begin{aligned} P(X < 0.8) &= N_{1,0.25^2}(0.8) = N_{0,1}(\frac{0.8-1.0}{0.25}) = N_{0,1}(-0.8) = 1 - N_{0,1}(0.8) = 0.2119. \\ P(Y < 0.8) &= N_{1.1,0.35^2}(0.8) = N_{0,1}(\frac{0.8-1.1}{0.35}) = N_{0,1}(-0.86) = 1 - N_{0,1}(0.86) = 0.1949. \end{aligned}$$

(c) What is the probability for a tree on the Able Farm to grow between 0.7m and 0.9m in a year? what if the tree was on the Baker Farm?

Using the same random variables X and Y as above:

$$\begin{aligned} P(0.7 < X < 0.9) &= N_{1,0.25^2}(0.9) - N_{1,0.25^2}(0.7) = N_{0,1}(-0.4) - N_{0,1}(-1.2) = \\ &= 1 - N_{0,1}(0.4) - 1 + N_{0,1}(1.2) = N_{0,1}(1.2) - N_{0,1}(0.4) = 0.8849 - 0.6554 = 0.2295. \\ P(0.7 < Y < 0.9) &= N_{1.1,0.35^2}(0.9) - N_{1.1,0.35^2}(0.7) = N_{0,1}(-0.57) - N_{0,1}(-1.14) = \\ &= N_{0,1}(1.14) - N_{0,1}(0.57) = 0.8729 - 0.7157 = 0.1572. \end{aligned}$$

(d) Assume, you've got two trees. One from Able and one from Baker. What can you say about the difference D in their growth of heights after one year? What is the distribution of D?

The difference between the growth of heights of two trees after one year D can be written as:

$$D = X - Y$$

From the reproductive property of normal distributions we know, that D, again, has a normal distribution; $D \sim N_{\mu_D,\sigma_D^2}$. We find the values for the parameters by looking at the parameters of X and Y:

$$\begin{aligned} \mu_D &= E[D] = E[X-Y] = E[X] - E[Y] = 1 - 1.1 = -0.1 \\ \sigma_D^2 &= Var[D] = Var[X + (-1)\cdot Y] = Var[X] + (-1)^2 Var[Y] = 0.25^2 + 0.35^2 = 0.185. \end{aligned}$$

Therefore, $D \sim N(-0.1, 0.43^2)$.

(e) On average, trees from the Baker farm will grow more than Able's trees. But what is the exact probability that a Baker tree has grown more than an Able tree in one year?

Using the difference in heights from the previous problem, the probability that a tree from the Baker farm is higher than a tree from the Able farm is equal to the probability that D is less than 0:

$$P(D < 0) = N_{-0.1,0.43^2}(0) = N_{0,1}(\frac{0.1}{0.43}) = N_{0,1}(0.23) = 0.5910.$$

2 Two Dimensional Density

Suppose that X and Y have a joint probability density function

$$f(x,y) = \begin{cases} k(x+2y) & \text{if } (x,y) \in [0.1] \times [0,1] \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find k such that $f(x,y)$ actually IS a density function.

We have two conditions for density functions: (1) $f_{X,Y}(x,y) \geq 0$ for all (x,y) and (2) $\int\int f_{X,Y}(x,y)dxdy = 1$.

For $k > 0$ the first condition is true, we need to make sure that the second also holds:

$$\begin{aligned} 1 &= \int_x \int_y k(x+2y)dydx = k\int_0^1 xy + y^2 \mid_0^1 dx = k\int_0^1 x + 1dx = k \cdot (0.5x^2 + x) \mid_0^1 = k \cdot 1.5 \\ &\Rightarrow k = \frac{2}{3}. \end{aligned}$$

(b) Find the marginal density of Y.

$$f_Y(y) = \int_0^1 \frac{2}{3}(x+2y)dx = \frac{2}{3}(0.5x^2 + 2xy) \mid_0^1 = \frac{1}{3}(1+4y)$$

(c) Find $E[Y]$.

$$E[Y] = \int_0^1 y \cdot \frac{1}{3}(1+4y)dy = \frac{1}{3}(\frac{1}{2}y^2 + \frac{4}{3}y^3)\Big|_0^1 = \frac{1}{6} + \frac{4}{9} = \frac{11}{18} \approx 0.61$$

(d) Find the density of $Z = Y^2$

Instead of looking for the density function directly. we have a look at Z's distribution function F_Z first. From the distribution function we can get the density function by differentiating.

$$\begin{aligned} F_Z(z) &= P(Z \leq z) = P(Y^2 \leq z) = P(-\sqrt{z} \leq Y \leq \sqrt{z}) = F_Y(\sqrt{z}) - F_Y(-\sqrt{z}) = \\ &= F_Y(\sqrt{z}) - 0. \end{aligned}$$

Since we don't know F_Y, we can differentiate the above expression using the chain rule (then we don't need to know F_Y):

$$f_Z(z) = F'_Z(z) = f_Y(\sqrt{z}) \cdot \frac{1}{2\sqrt{z}} = \frac{1}{3}(1 + 4\sqrt{z})\frac{1}{2\sqrt{z}} = \frac{1}{6}(\frac{1}{\sqrt{z}} + 4)$$

(e) Find the conditional density $f_{X|Y}(x \mid y)$.

$$\begin{aligned} f_{X|Y}(x \mid y) &= \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{\frac{2}{3}(x + 2y)}{\frac{1}{3}(1 + 4y)} \textit{ for } x, y \in [0,1] = \\ &= \frac{2x + 4y}{1 + 4y} \textit{ for } x, y \in [0,1] \end{aligned}$$

(f) Find $P(X \geq 0.5 \mid Y = 0.5)$.

$$\begin{aligned} P(X \geq 0.5 \mid Y = 0.5) &= \int_{0.5}^{1} f_{X|Y}(x \mid 0.5)dx = \int_{0.5}^{1} \frac{2x + 4 \cdot 0.5}{1 + 4 \cdot 0.5}dx = \frac{1}{3}\int_{0.5}^{1} 2x + 2dx = \\ &= \frac{1}{3}(x^2 + 2x) \mid_{0.5}^{1} = \frac{1}{3}(3 - \frac{1}{4} - 1) = \frac{7}{12}. \end{aligned}$$

3 Grand Central Station

Two friends A and B meet every morning at the Grand Central Station around 7 am. Suppose the actual times they arrive are independent and uniformly distributed between 6:55 am and 7:05am. Let Z denote the time between arrivals, i.e. Z = time B arrives - time A arrives (can be negative!).

(a) What is the range of Z?

$Z := B - A$ *is a continuous random variable with values in* $[-10, 10]$.

(b) Find the density of Z.

A and B are uniform random variables on the 10 min interval $[6 : 55, 7 : 05]$. *To make the math simpler, we assume, we are dealing with the interval* $[0, 10]$.

Then the density functions f_A and f_B are

$$f_A(a) = f_B(b) = \frac{1}{10} \textit{ for } a, b \in [0, 10]$$

In order to find the density of Z, we need to use the convolution theorem.

$$f_Z(z) = \int_0^{10} f_B(w) f_A(w+z) dw$$

We know: for $0 \leq w \leq 10$ $f_B(w) = 0.1$*, for* $0 \leq z + w \leq 10$ $f_A(z+w) = 0.1$. *The second condition is equivalent to* $-z \leq w \leq 10 - z$.

Combining these conditions gives:

$f_B(w) f_A(z+w) = 0.01$ *for* $\max(0, -z) \leq w \leq \min(10, 10-z)$

Therefore

$$\begin{aligned} f_Z(z) &= \int_{\max(0,-z)}^{\min(10,10-z)} 0.01 dw = 0.01(\min(10, 10-z) - \max(0,-z)) = \\ &= 0.01 \cdot \begin{cases} 10 + z & \textit{for } -10 \leq z < 0 \\ 10 - z & \textit{for } 0 \leq z \leq 10 \end{cases} \end{aligned}$$

Z has a symmetric tent shaped density function f_Z with peak in 0.

(c) Find the probability that A waits for at least 5 minutes before B arrives.

The event that A waits for 5 minutes before B arrives corresponds to $Z > 5$. Geometrically, we can find the probability $P(Z > 5)$ as the area under the density function. From above, we know that this area is triangular, with height 0.05 and width 5; therefore $P(Z > 5) = 0.125$.

Using the density function in analytically, we compute

$$P(Z > 5) = \int_5^{10} f_Z(z) dz = 0.01 \cdot \int_5^{10} 10 - z dz = 0.01(10z - 0.5z^2) \mid_5^{10} = 0.125.$$