Homework5

Laurenz Ohnemüller

10/7/2019

Question 1

Volume of the unit ball: $v_n = \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2}+1)}$
Gamma function: $\Gamma(a) = \int_0^\infty x^a e^{-x} \frac{dx}{x}$

(a)
First of all, we want to compute the probability that a point is in the unit ball in the n-dimensional space. Let X denote that $(U_1, U_2, ..., U_n)$ is in the unit ball in R^n then:

$$P(X_n) = \frac{volume\_of\_ball}{area\_of\_square}$$

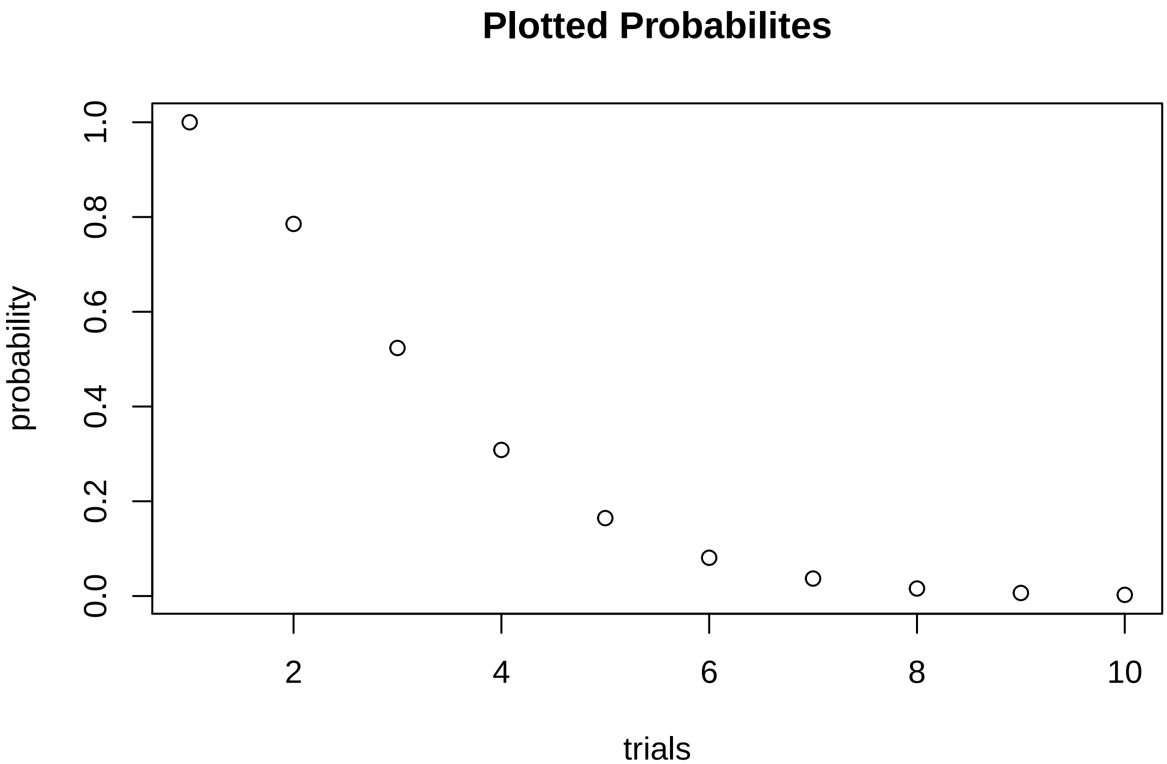
We can calculate this by dividing the volume of the unit ball v_n by the area of the square of dimension n. The volume for the unit ball in the nth dimension is already given, so we only need to figure out the area of the square.
Since $(U_1, U_2, ..., U_n) \sim Unif(-1, 1)$, we can imagine the square going from -1 to 1, which means that one side has the length of 2. Now we can just take it to the power of n, to find the area of the nth dimension. So:

$$P(X_n) = \frac{v_n}{2^n} = \frac{\frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2}+1)}}{2^n} = \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2}+1)2^n}$$

Finally, we just have to insert a specific dimension and we get the probability that this particular point is in the unit ball.

(b)

```r
# I assumed we are allowed to use this library, since we also used it multiple times in the Exercises
library("tibble")
trials = 10
p = tibble(n=1:trials,
           (pi^(n/2))/(gamma(n/2+1)*2^n)
)
plot(p, ylab = "probability", xlab = "trials", main="Plotted Probabilites")
```



(c) We can see the condition $|U_j| > c$ as a Bernoulli experiment, either this is true ("success") or it is not true ("failure"). Having n of these Bernoulli trials would speak for a Binomial distribution, and to that we also are sure about the independence of U_j, which insures us that the probability p does not change. So:

$$X_n \sim Bin(n, p)$$

$$X_n = \binom{n}{k} p^k (1-p)^{n-k}$$

(d) Imagine a circle drawn inside a square - to make it more simple, we can do that in the third dimension, so with a ball inside a cube. We now that each edge has the length 2, since one goes from -1 to 1, while the ball also surrounds the origin (0,0). We want to find the probability that a random point (U_1, U_2, U_3) is inside this unit ball.

Question 2

The definition of the PMF of a joint distribution is:

$$p_{x,y} = P(X = x, Y = y)$$

(a) In this question, we are supposed to find the joint PMF of X, Y and N.

$$P(X = x, Y = y, N = n)$$

Note that x, y and n are nonnegative integers. Also note that since x describes the number of times the traveler is lost and asks for direction and y the number of times he does not ask for direction, while n describes the amount of times he is lost in generel. This sum $x+y = n$ must hold, otherwise the PMF is 0.

Now we apply the Law of Total Probability and condition on N:

$$P(X=x,Y=y,N=n)=\sum_{n=0}^{\infty}P(X=x,Y=y|N=n)P(N=n)=P(X=x,Y=y|N=n)P(N=n)$$

Hereby, we can take the N out, since we know that it holds in the case where $x+y=n$, which means we can replace it by $N=X+Y$.

$$P(X=x,Y=y,N=n)=P(X=x,Y=y,X+Y=n)=P(X=x,Y=y,X=n-Y)=P(X=x,Y=y,X=n-y)$$

We can observe that $X=n-y$ is redundant information, since it the exact same thing as $X=x$, so we end up here:

$$P(X=x,Y=y)$$

From the chicken and egg story (and also proven in (c)), we know that X and Y are independet and Possion distributed:
$X \sim Pois(\lambda p)$
$Y \sim Pois(\lambda q)$, whereby q = 1 - p
, which brings us to this:

$$P(X=x,Y=y)=P(X=x)P(Y=y)=\frac{e^{-\lambda p}(\lambda p)^x}{x!}\frac{e^{-\lambda q}(\lambda q)^y}{y!}$$

Lastly, we want to check wheather X, Y and are independent or not. For independence this condition must hold: $P(X=x,Y=y,N=n)=P(X=x)P(Y=y)P(N=n)$ We can already see that this will not hold, since we just take the calculated value from above and multiply it bei $P(N=n)$, which is also poisson distributed $N \sim Pois(\lambda)$ and it won't equal each other. Summarizing, (X,Y,N) are dependent.

(b)

In the second part we want to find the joint PMF of X and N, which we can express in the following way using again the Law of Total Probability:

$$P(X=x,N=n)=P(X=x|N=n)P(N=n)$$

Here, we plug in the defintions of the Binomial (for $P(X=x|N=n)$)and Possion distribution (for $P(N=n)$):

$$=\binom{n}{x}p^x q^{n-x}*\frac{e^{-\lambda}\lambda^n}{n!}$$

To check wheather they are dependent or independent, we use this condition of independence:

$$P(X=x,N=n)=P(X=x)P(N=n)$$

We first calculate $P(X=x)$, where the random variable has the folllowing distribution $X \sim Pois(\lambda p)$:

$$P(X=x)=\frac{e^{-\lambda p}*(\lambda p)^x}{x!}$$

and since $P(N=n)$ is normally Poission distributed, we get:

$$P(X=x)P(N=n)=\frac{e^{-\lambda p}*(\lambda p)^x}{x!}*\frac{e^{-\lambda}*\lambda^n}{n!}\neq P(X=x,N=n)$$

This means that N and X are indeed dependent.

(c)

Finally, we are supposed to find:

$$P(X = x, Y = y)$$

, where we condition on N and apply the Law of total probability (LOTP), we assume we know the amount of times our traveler was lost.

$$P(X = x, Y = y) = \sum_{n=0}^{\infty} P(X = x, Y = y|N = n)P(N = n)$$

Since n describes the amount of times the traveler gets lost, while x, as well as y, describe if he asked for directions or not in the case that he is lost, we can say $x + y = n$, otherwise the probability is zero.
Knowing this, we can transform our equation to:

$$P(X = x, Y = y|N = x + y)P(N = x + y)$$

Now we can observe that that $Y = y$ is redundant, since it is exactly the same event than $X = x$.
So we are left with:

$$P(X = x, |N = x + y)P(N = x + y)$$

, which is easy to compute using the definitions of the Binomial and Possion distribution.

Binomial distribution: $\frac{n!}{k!(n-k)!}p^k q^{n-k}$, where n = x+y, and k = x

Possion distribution: $\frac{e^{-\lambda}\lambda^k}{k!}$, where k = x+y

$$= \frac{(x + y)!}{x!y!} p^x q^y * \frac{e^{-\lambda}\lambda^{x+y}}{(x + y)!}$$

Here we can reduce and reformate the equation, sucht that:

$$= e^{-\lambda p}\frac{(\lambda p)^x}{x!} * e^{-\lambda q}\frac{(\lambda q)^y}{y!}$$

On this equation, we can observe that X and Y are independent, since the joint PMF is the product of $X \sim Pois(\lambda p)$ and $X \sim Pois(\lambda q)$, which makes this condition hold: $P(X = x, Y = y) = P(X = x)P(Y = y)$.

Question 3

The joint PDF of X and Y:

$$f_{X,Y}(x, y) = ce^{-\frac{x^2}{2}}e^{-\frac{y^2}{2}}$$

(a)

To be a valid PDF $f_{X,Y}(x, y)$ must satisfy two criteria:

(1) Nonnegative: $f_{X,Y}(x, y) \geq 0$;

(2) Integrates to 1: $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x, y)dxdy = 1$

This allows us directly to define c as $c \geq 0$, since otherwise (1) would be violated. To find the actual c, we have to solve the integral:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} ce^{-\frac{x^2}{2}}e^{-\frac{y^2}{2}}dxdy = 1$$

c is constant, which means that we can pull it before the integrals:

$$c\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}}e^{-\frac{y^2}{2}}dxdy = 1$$

Now, we will just focus on solving this integral - without thinking about the 1 or the c:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{x^2+y^2}{2}}dxdy$$

Here we make use of the polar coordinates (A.7.2 in appendix in our Probability book).
$x = r * cos(\theta), y = r * sin(\theta)$, where r is the distance from (x,y) to the origin and $\theta \in [0, 2\pi]$ is the angle. Hereby, $dxdy$ becomes $rdrd\theta$ and $r = \sqrt{x^2 + y^2}$ (so r must be $r \geq 0$). This changes also the limits, as already explained θ goes from 0 to 2π, while r can take on the values form 0 to ∞ and it follows:

$$\int_{0}^{2\pi}\int_{0}^{\infty} e^{-\frac{r^2}{2}}rdrd\theta$$

This integral is now solvable, using substitution $u = \frac{r^2}{2}; du = rdr$:

$$\int_{0}^{2\pi}(\int_{0}^{\infty} e^{-u}du)d\theta = \int_{0}^{2\pi}(-e^{-u}|_{0}^{\infty})d\theta = \int_{0}^{2\pi} 1d\theta = 2\pi \tag{1}$$

Finally, we can plug this result into our equation from earlier:

$$c * 2\pi = 1 \rightarrow c = \frac{1}{2\pi}$$

(b)

Since our distribution is continous, we need the definition of the Marginal PDF of X:

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy$$

From here on we can continue our work similarly to the previous question:

$$f_X(x) = \int_{-\infty}^{\infty} ce^{-\frac{x^2}{2}}e^{-\frac{y^2}{2}}dy = ce^{-\frac{x^2}{2}}\int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}dy$$

To figure out $\int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}dy$, we can just look at (a) again, and in particular we can focus on equation (1), where we proved that:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}}e^{-\frac{y^2}{2}}dxdy = 2\pi$$

We can split that integral apart and summarize it as, since x and y are just different names for the "dummy" variable z:

$$\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy = (\int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy)^2 = 2\pi$$

Out of this equality, we find the value for our integral:

$$\int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy = \sqrt{2\pi}$$

Now lets put everything together again and plug in the values for the integral and c:

$$f_X(x) = ce^{-\frac{x^2}{2}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy = \frac{1}{2\pi} e^{-\frac{x^2}{2}} * \sqrt{2\pi} = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

On the exact same way, we can also find out $f_Y(y)$, which will be - because of symmetry - the same as $f_Y(y)$:

$$f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$$

This is by the way the standard normal distribution.
To check wheather the X and Y are independent, we examine this condition:

$$f_{X,Y}(x, y) = f_X(x) f_Y(y)$$

Let us now insert our solutions and check the condition:

$$\frac{1}{2\pi} e^{-\frac{x^2}{2}} e^{-\frac{y^2}{2}} \stackrel{?}{=} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} * \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$$

This is actually true, which confirms the condition and makes the X and Y independent.
(c)
First of all, we know that X and Y are both marginally $\mathcal{N}(0, 1)$. For (X, Y) to be a Bivariate Normal, $aX + bY$ has to have a normal distribution for $a, b \in \mathbb{R}$.
This is certainly true,the sum of independent Normals is also Normal. Any value for a or b, won't change that. Also for $a = b = 0$, where $aX + bY = 0$ is still normally distributed with mean and variance 0.