Tutorial Week 8 – Continuous Distributions MATH 4043

Using **R** to calculate the integrals.

1. Integrate the following integrals.

   (a) $\int_0^\infty 5xe^{-5x}dx$.

   (b) $\int_0^\infty 5xe^{-5x^2}dx$.

   (c) $\int_1^2 x^{-1}\ln(x)dx$.

   Use **R** to do integration of definite integrals. For example for $\int_0^2 x^2dx$, it is done this way

```
> integrand <- function(x) {x^2}
> integrate(integrand,0,1)
0.3333333 with absolute error < 3.7e-15
```

   So you can use it to check your results. Use Inf for ∞.

2. The number of visitors clicking an online shopping website is distributed as Poisson with a mean visit rate of 7 per hour. Any individual visitor who goes onto the website has a 65% chance of making a purchase. Assume each individual visitor acts independently of each other, and assume that the visitor only makes 1 click to the website within that hour.

   Let N be the number of visits to the site. Let X be the number of purchases made.

   (a) What is the probability that within a particular hour, there are 5 visits to the site?

   (b) What is the probability that given n visits to the site, there are x purchases where $x = 0, 1, 2, \cdots, n$?

   (c) What is the probability that there are 5 visits to the site with only 3 purchases?

   (d) Simulate using **R** the number of actual purchases over 10 hour period.

3. Refer to the Question 2 above
   The number of visitors clicking an online shopping website is distributed as Poisson with a mean visit rate of 7 per hour. Any individual visitor who goes onto the website has a 65% chance of making a purchase. Assume each individual visitor acts independently of each other, and assume that the visitor only makes 1 click to the website within that hour. If the ith visitor makes a purchase, the amount spent Y_i is independently distributed as exponential Exp(λ_e) with $\lambda_e = 0.05$. Simulate the number of visitors in a

two hour period, the number from these visitors who make a purchase, and if they make a purchase, their purchase amounts.

4. Refer to Q2 above.
   Suppose now that if the ith visitor makes a purchase, the log of the amount spent $\ln(Y_i)$ is normally distributed as $N(3, 2)$. Simulate the number of visitors in a two hour period, the number from these visitors who make a purchase, and if they make a purchase, their purchase amounts.

Solutions

1. Integrate the following integrals.

   Solution.

   (a) $\int_0^\infty 5xe^{-5x}dx$.
   We rewrite this as $\int_0^\infty x \times 5e^{-5x}dx$. We apply integration by parts with $u = x$, $dv = 5e^{-5x}dx$, $du = dx$ and $v = -e^{-5x}$. Thus

$$\begin{aligned}
\int_0^\infty x \times 5e^{-5x}dx &= -xe^{-5x}|_0^\infty - \int_0^\infty (-e^{-5x})dx \\
&= -0 - (-0) + \int_0^\infty e^{-5x}dx \\
&= -\left.\frac{e^{-5x}}{5}\right|_0^\infty \\
&= -0 - \left(-\frac{e^0}{5}\right) = \frac{1}{5}.
\end{aligned}$$

   Using **R** , we can try

```
> integrand <- function(x) {5*x*exp(-5*x)}
> integrate(integrand,0,Inf)
0.2 with absolute error < 8.3e-05
```

   (b) $\int_0^\infty 5xe^{-5x^2}dx$.
   We can do by substitution $u = 5x^2$, so that $du = 10xdx$ and $dx = \frac{du}{10x}$. Note that we can also substitute e.g. $u = -5x^2$ or even just $u = x^2$ and so on, but the interval limits for the integral for u must be consistent with the corresponding interval limits for x.

$$\begin{aligned}
\int_0^\infty 5xe^{-5x^2}dx &= \int_0^\infty 5xe^{-u}\frac{du}{10x} \\
&= \frac{1}{2}\int_0^\infty e^{-u}du \\
&= -\left.\frac{e^{-u}}{2}\right|_0^\infty \\
&= -0 - \left(-\frac{e^0}{2}\right) = \frac{1}{2}.
\end{aligned}$$

```
> integrand <- function(x) {5*x*exp(-5*x^2)}
> integrate(integrand,0,Inf)
0.5 with absolute error < 1.7e-05
```

(c) $\int_1^2 x^{-1}\ln(x)dx$.

We rewrite as $\int_1^2 \ln(x)[x^{-1}]dx$. We apply integration by parts with $u = \ln(x)$, $dv = x^{-1}dx$, $du = x^{-1}dx$ and $v = \ln(x)$. Thus

$$\int_1^2 \ln(x)[x^{-1}]dx = (\ln(x))^2|_1^2 - \int_1^2 [\ln(x)]x^{-1}dx$$

$$\text{So } \int_1^2 \ln(x)[x^{-1}]dx + \int_1^2 \ln(x)[x^{-1}]dx = (\ln 2)^2 - (0)^2$$

$$\text{Thus } \int_1^2 x^{-1}\ln(x)dx = \frac{(\ln 2)^2}{2}.$$

```
> integrand <- function(x) {x^(-1)*log(x)}
> integrate(integrand,1,2)
0.2402265 with absolute error < 2.7e-15
> 0.5*(log(2))^2
[1] 0.2402265
```

2. The number of visitors clicking an online shopping website is distributed as Poisson with a mean visit rate of 7 per hour. Any individual visitor who goes onto the website has a 65% chance of making a purchase. Assume each individual visitor acts independently of each other, and assume that the visitor only makes 1 click to the website within that hour.

Solution. Let N be the number of visits to the site. Let X be the number of purchases made.

(a) What is the probability that within a particular hour, there are 5 visits to the site?
$N \sim \text{Pois}(7)$, thus

$$P[N = 5] = \frac{e^{-7}7^5}{5!} = 0.1277.$$

(b) What is the probability that given n visits to the site, there are x purchases where $x = 0, 1, 2, \cdots, n$?
Basically this question asks us to find the conditional probability $P[X = x|N = n]$ where $x = 0, 1, \cdots, n$. Note that given $N = n$, there are n independent visitors to the site. So this is like a binomial problem where if an individual visitor makes a purchase, it is like getting heads for a coin toss. Therefore the distribution of $X|N = n$ is actually $\text{Bin}(n, 0.65)$, since each visitor has a 0.65 probability of making a purchase. So we have

$$P[X = x|N = n] = C_x^n(0.65)^x(0.35)^{n-x},$$

for $x = 0, 1, \cdots, n$ by applying the formula for the binomial probability mass function.

(c) What is the probability that there are 5 visits to the site with only 3 purchases?

Applying the formula from part (b), this is

$$\begin{aligned} P[\{X=3\} \cap \{N=5\}] &= P[X=3|N=5]P[N=5] \\ &= C_3^5(0.65)^3(0.35)^2 \times 0.1277 \\ &= 0.3364 \times 0.1277 = 0.04296. \end{aligned}$$

(d) Simulate using **R** the number of actual purchases over 10 hour period.

First we simulate the number visits regardless of purchases made. Store the various values of the N obtained and then simulate a sample from each binomial with each of the N.

```
> N <- rpois(10,7)
> N
 [1]  9  9 10  5  4  3  6  9  6  4
```

Next we have to do individually with a binomial random sample of size 1 for each hour, e.g. repeat each step for 10 times and type in $n = 9, 9, 10, 5, 4, 3, 6, 9, 6, 4$.

```
> rbinom(1,9,0.65)
 [1] 7
> rbinom(1,9,0.65)
 [1] 5
 > rbinom(1,10,0.65)
 [1] 6
 ...
```

etc. and I obtained 7 5 6 4 2 3 4 6 5 1 for the corresponding number of purchases.

So to interpret the results, we have 9 visits and 7 purchases made during the 1st hour, 9 visits and 5 purchases made during the 2nd hour, etc.

Note that when I did rbinom(10,N,0.65), it also appears to give me 10 sensible results for X, the number of purchases right away. It looks like if N is entered as a vector object instead of single number n, **R** appears to simulate it as 10 samples of binomial with the differing n corresponding to the respective entries in N. However, I am not able to find concrete documentation that this is how **R** is interpreting it.

3. Refer to the Question 2 above

The number of visitors clicking an online shopping website is distributed as Poisson with a mean visit rate of 7 per hour. Any individual visitor who goes onto the website has a 65% chance of making a purchase. Assume each individual visitor acts independently of each other, and assume that the visitor only makes 1 click to the website within that hour. If the ith visitor makes a purchase, the amount spent Y_i is independently distributed as exponential Exp(λ_e) with $\lambda_e = 0.05$. Simulate the number of visitors in a two hour period, the number from these visitors who make a purchase, and if they make a purchase, their purchase amounts.

Solution. Let N be the number of visitors and X be the number of visitors that make a purchase. Note that there are several ways in which you can set up or organise your simulation.

```
> N <- rpois(1,7)
> N
[1] 4
> X <- rbinom(N,1,0.65)
> X
[1] 0 1 0 1
> Y <- rgamma(N,1,0.05)
> Y
[1]  4.448685  8.826599  3.917239 27.486900
> X*Y
[1]  0.000000  8.826599  0.000000 27.486900
> sum(X*Y)
[1] 36.3135
```

In my case, I have 4 visits, 2 of which end up with a purchase, namely the 2nd and 4th visitors. $X * Y$ blanks out the dollar amounts of the 4.448685 and 3.917239, since they are not needed as the 1st and 3rd customers did not make any purchases. The total amount spent by the visitors is \$33.3135.

4. Refer to Q2 above.
Suppose now that if the ith visitor makes a purchase, the log of the amount spent $\ln(Y_i)$ is normally distributed as $N(3, 2)$. Simulate the number of visitors in a two hour period, the number from these visitors who make a purchase, and if they make a purchase, their purchase amounts.
Solution.

```
> N <- rpois(1,7)
> N
[1] 7
> X <- rbinom(N,1,0.65)
> X
[1] 1 1 1 0 0 1 1
> Y <- rnorm(N,3,2)
> Y
[1] 4.974824 5.449090 5.830086 1.957981 6.927325 3.700344 3.171369
```

```
> Y <- exp(Y)
> Y
[1]  144.72335  232.54640  340.38795    7.08501 1019.76255   40.46122   23.84010
> X*Y
[1] 144.72335 232.54640 340.38795   0.00000   0.00000  40.46122  23.84010
> sum(X*Y)
[1] 781.959
```

In my case, I have 7 visits, 5 of which end up with a purchase. $X * Y$ blanks out the dollar amounts of those visitors who did not make any purchases. The total amount spent by the visitors is \$781.959.