Problems on continuous random variables

Ignacio Cascos, Department of Statistics, Universidad Carlos III de Madrid

2018/19

1. Consider $X \sim \mathrm{N}(\mu = 1, \sigma = 2)$, $Y \sim \mathrm{Exp}(\lambda = 3)$, $U \sim \mathrm{U}(2, 5)$, $V \sim \mathrm{B}(n = 10, p = 0.3)$, and $W \sim \mathcal{P}(\lambda = 5)$ Compute the following:
    1. $P(X \geq 4)$.
    2. $P(3 \leq X \leq 5)$.
    3. $P(|X| \geq 1)$.
    4. $P(X > 2 | X > 1)$.
    5. Determine x such that $P(X > x) = 0.35$.
    6. $P(Y > 5)$.
    7. $P(Y < 5 | Y > 2)$.
    8. Determine y such that $P(Y < y) = 0.15$.
    9. $P(U > 3.5 | U > 1)$.
    10. $P(V \geq 7)$.
    11. Determine the smallest v such that $P(V > v) \leq 0.05$.
    12. $P(W > 5)$.

```
1-pnorm(4,mean=1,sd=2)
```

```
## [1] 0.0668072
```

```
pnorm(5,mean=1,sd=2)-pnorm(3,mean=1,sd=2)
```

```
## [1] 0.1359051
```

```
pnorm(-1,mean=1,sd=2)+(1-pnorm(1,mean=1,sd=2))
```

```
## [1] 0.6586553
```

```
(1-pnorm(2,mean=1,sd=2))/(1-pnorm(1,mean=1,sd=2))
```

```
## [1] 0.6170751
```

```
qnorm(0.65,mean=1,sd=2)
```

```
## [1] 1.770641
```

```
1-pexp(5,rate=3)
```

```
## [1] 3.059023e-07
```

```
(pexp(5,rate=3)-pexp(2,rate=3))/(1-pexp(2,rate=3))
```

```
## [1] 0.9998766
```

```
qexp(.15,rate=3)
```

```
## [1] 0.05417298
```

```
(1-punif(3.5,min=2,max=5))/(1-punif(1,min=2,max=5))
```

```
## [1] 0.5
```

```
1-pbinom(6,size=10,prob=0.3)
```

```
## [1] 0.01059208
```

```r
qbinom(0.95,size=10,prob=0.3)
```

```
## [1] 5
```

```r
1-ppois(5,lambda=5)
```

```
## [1] 0.3840393
```

2. Potholes on a highway can be a serious problem, and are in constant need of repair. With particular types of terrain and asphalt, past experience suggests that there are, on average, 2 potholes per mile after a certain amount of usage, and the random variable 'number of potholes' can be modeled by means of a Poisson distribution.
   1. What is the probability that more than 25 potholes appear in a section of 10 miles? Denote by X the number of potholes in 10 miles. Clearly X follows a Poisson distribucion with parameter $\lambda = 20$ potholes per 10 miles, $X \sim \mathcal{P}(\lambda = 20)$ and $P(X > 25)$ is

   ```r
   1-ppois(25,lambda=20)
   ```

   ```
   ## [1] 0.112185
   ```

   2. A group of maintenance workers detected the first pothole of the day at a distance of half a mile from the point where they started working. How many more miles should they inspect in order to detect another pothole with probability 0.9? Denote by T the distance (in miles) between two consecutive potholes, then $T \sim \text{Exp}(\lambda = 2)$, and we want to determine the distance in miles that should be inspected t such that $P(T < t + 1/2 | T > 1/2) = 0.9$, so $P(T < t) = 0.9$,

   ```r
   qexp(0.9,rate=2)
   ```

   ```
   ## [1] 1.151293
   ```

3. The compressive strength of samples of cement can be modeled by a normal distribution with a mean of 6000 kilograms per square centimeter and a standard deviation of 100 kilograms per square centimeter.
   1. What is the probability that a sample's strength is less than 6250 kg/cm^2?

   ```r
   pnorm(6250,mean=6000,sd=100)
   ```

   ```
   ## [1] 0.9937903
   ```

   2. What is the probability that a samples's strength is between 5800 and 5900 kg/cm^2?

   ```r
   pnorm(5900,mean=6000,sd=100)-pnorm(5800,mean=6000,sd=100)
   ```

   ```
   ## [1] 0.1359051
   ```

   3. What strength is exceeded by 95% of the samples?

   ```r
   qnorm(.05,mean=6000,sd=100)
   ```

   ```
   ## [1] 5835.515
   ```

   4. Only samples whose strength is greater than 5750 kg/cm^2 are acceptable. What is the probability that in a group of 10 samples, at least 9 of them are acceptable? Denote by p the probability that the strength of a sample is greater than 5750. The number of samples (out of 10) whose strength is greater than 5750 follows a Binomial distribution with parameters $n = 10$ and p, so

```r
    p=1-pnorm(5750,mean=6000,sd=100)
    1-pbinom(8,size=10,prob=p)
```

```
## [1] 0.9983213
```

4. The time in days, T, that a manufacturing system is out of operation every time it breaks down has cumulative distribution function

$$F_T(t) = \begin{cases} 1 - (2/t)^3 & \text{if } t > 2 \\ 0 & \text{otherwise} \end{cases} .$$

   Apply the inverse transform method in order to simulate from the cdf above and provide the exact answer together with an approximate answer to each of the following questions. The approximate answers are to be obtained after 10000 simulations (use NIU, DNI, or NIE as seed).

1. What is the probability that the system is out of operation longer than 3 days but shorter than 5? If the system has been out of operation for 3 days, what is the probability that it is out of operation less than 5 days?
2. What out of operation time is exceeded 75% of the times the system breaks down?
3. Compute the mean and variance of the out of operation time.
4. If the breakdown cost for the company is $Y = T^2$, simulate from it and compute its mean. We start with the exact answers:

- $P(3 < T < 5) = F_T(5) - F_T(3) = 0.2323$.
- $P(T < 5|T > 3) = P(3 < T < 5)/P(T > 3) = 0.2323/0.2963 = 0.784$.
- $P(T > t) = 0.75$, so $(2/t)^3 = 0.75$ and $t = 2.2013$.

$$f_T(t) = F_T'(t) = \begin{cases} 24t^{-4} & \text{if } t > 2 \\ 0 & \text{otherwise} \end{cases}.$$

- $\mathbb{E}[T] = \int_{-\infty}^{\infty} x f_T(x)dx = \int_2^{\infty} 24x^{-3}dx = [-12x^{-2}]_2^{\infty} = 3$.
- $\mathbb{E}[T^2] = \int_{-\infty}^{\infty} x^2 f_T(x)dx = \int_2^{\infty} 24x^{-2}dx = [-24x^{-1}]_2^{\infty} = 12$.
- $\text{Var}[T] = \mathbb{E}[T^2] - \text{E}[T]^2 = 12 - 3^2 = 3$.
- The quantile funciton of T (inverse cdf) is given by the soluction on t to the equation $1-(2/t)^3 = p$ which is $F_T^{-1}(p) = 2(1-p)^{-1/3}$ for $0 < p < 1$. We use it in our simulations.

```
n=100000
set.seed(2)
simulunif=runif(n)
simulF=2*(1-simulunif)^(-1/3)
sum((simulF>3)*(simulF<5))/n
```

```
## [1] 0.23365
```

```
sum((simulF>3)*(simulF<5))/sum(simulF>3)
```

```
## [1] 0.7810724
```

```
quantile(simulF,0.25)
```

```
##      25%
## 2.200841
```

```
mean(simulF)
```

```
## [1] 3.008873
```

```
var(simulF)
```

```
## [1] 2.880873
```

```
mean(simulF^2)
```

```
## [1] 11.93416
```

5. Use the normal approximation to the Binomial distribution to solve the following problem. Four flatmates Alice, Bob, Charly, and Dave roll a 6-sided die every night in order to decide who washes the dishes after dinner. If the outcome is 1, Alice washes, it is its 2, Bob does, while for 3 or 4 Charly must wash the dishes and for 5 or 6 it is Dave's turn.
   1. How many days (dinners) must we wait until Bob washes the dishes at least 11 times with probability 0.95? If we set $X \equiv$ 'number of dinners Bob washes the dishes in n days' then $X \sim \text{B}(n, p = 1/6)$ with n unknown, and we have to determine n such that $P(X \geq 11) \geq 0.95$. After the DeMoivre-Laplace Theorem, $X \approx \text{N}(np, \sqrt{np(1-p)})$, so applying a continuity correction and standardizing $P(X \geq 11) = P(X \geq 10.5) = P(Z \geq (10.5 - np)/\sqrt{np(1-p)}) \geq 0.95$, where $Z \sim \text{N}(0, 1)$. Finally $(10.5 - np)/\sqrt{np(1-p)}$ is the 0.05 quantile of a standard normal random variable, `qnorm(0.05)`.

In conclusion $\sqrt{n} = 9.987$ and $n = 99.74$. Clearly we need that n assumes an integer value, so we can guarantee that Bob washes at least 11 nights with probability (at least) 0.95 in $n = 100$ nights.

2. What is the probability that we have to wait longer than that time until Bob washes the dishes at least 11 times? For the previous value of n compute $P(X < 11)$ using either the Binomial distribution or its normal approximation.

```
pbinom(10,size=100,prob=1/6)
```

```
## [1] 0.04269568
```

```
pnorm(10.5,mean=100/6,sd=sqrt(100*5/36))
```

```
## [1] 0.04899367
```