Вероятности и статистика с R

Асен Чорбаджиев

April 11, 2017

1 Генератори

Равновероятно разпределени стойности в интервала [a,b] се описват с непрекъснато равномерно разпределение. Функцията в R е unif(), със съответните представки d,p,q,r за вероятност, разпределение, квантил и генератор на случайни числа:
Пример: Да се генерират 100 случайни равномерно разпределени интервала от [0,2] и да се начертае хистограма и теоретична плътност.

```
x=runif(100,0,2)
hist(x,probability=TRUE,col=gray(.9),main="uniform on [0,2]")
curve(dunif(x,0,2),add=T)
```

Случаен избор от извадка се прави с функцията sample(). Параметърът за повторение е replace=TRUE.

```
sample(1:10,10)
sample(1:10,10,replace=TRUE)
```

2 Memoryless random distributions

Геометричното разпределение е частен случай на Отрицателно Биномното разпределение, когато броят на успешните биномни опити r=1. Това означава, че вероятността за първо случване на събитие с вероятност p след n опита е:

$$P(n) = p(1-p)^n \tag{1}$$

Функциите в R са с име geom(), със съответните представки d,p,q,r за вероятност, разпределение, квантил и генератор на случайни числа. Входни параметри са брой на предходни неуспехи и тяхната вероятност:

Пример: Брой опити до успех:
Y=0:10
plot(Y, dgeom(Y,0.6), type="h", ylim=c(0,1), main="Geometric distribution for p=0.6", ylab="P(Y=Y)", xlab="Y=Number of failures before first success")

Геометричното разпределение е дискретно memoryless random distribution. Когато процесите на continuous memoryless чакане до поява на събитие са непрекъснати, вероятностното разпределение е експоненциално. Това е еквивалентно на Поасоново разпределение с k=0:

$$P(x) = \frac{1}{\mu} e^{-x/\mu} \tag{2}$$

Функциите в R за експоненциално разпределение са exp, със съответните представки d,p,q,r за плътност, разпределение, квантил и генератор на случайни числа.

Пример: Среден живот на електрическа крушка е μ =2500 часа. Да се генерират 100 случайни експоненциални проби и да се начертае хистограма и теоретична плътност.
x=rexp(100,1/2500)
hist(x,probability=TRUE,col=gray(.9),main="exponential mean=2500")
curve(dexp(x,1/2500),add=T)

3 Нормално Разпределение

Нормалното разпределение разкрива специален клас симетрични и двупараметрични разпределение $N(\mu, \sigma^2)$:
(i) μ = Средно
(ii) σ = Стандартното отклонение на разпределението.
Функциите в R за нормално разпределение са norm(), със съответните представки d,p,q,r за плътност, разпределение, квантил и генератор на случайни числа.
Променяйки стойностите на μ и σ изместват графиката на разпределението:
x=c(1:200)
plot(dnorm(x, 100,10), ylim=c(0,0.1))
lines(dnorm(x, 100,5), col="red")
lines(dnorm(x, 130,10), col="blue")

Функцията на вероятностна плътност е:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} exp^{-(x-\mu)^2/(2\sigma^2)} \tag{3}$$

Стандартно Нормално разпределение наричаме Z=N(0,1). Параметрите по подразбиране на norm() са за Стандартно Нормално Разпределение. Формулата за нормализиране на Нормално разпределение е:

$$Z = \frac{X - \mu}{\sigma} \tag{4}$$

4 Упражнения:

- **Зад. 1.**: На тренировка баскетболист изпълнява наказателен удар до вкарване на кош. Начертайте разпределението на получаване на първи кош за p=1/3, p=1/2, p=2/3.
- **Зад. 2.**: Времето за ремонт и обслужване на автомобил след всеки транспортен курс е експоненциално със средно време 5 мин. Пресметнете вероятността времето за ремонт на поредният автомобил на стена да не превиши k мин. Пресметнете вероятността за k=30 min и начертайте графиката за k=1:30.
- **Зад. 3.**: Генерирайте 50 случайни стойнсти N(5,5). Нормализирайте ги и резултата сравнете с 50 случайни N(0,1)
- **Зад. 4.**: Ако Z = N(0, 1) на колко е равно P(Z > 0.92)? P(Z > -0.5)? P(-0.64 < Z < 0.43)?
- **Зад. 5.**: Ако X = N(3, 4), На колко е равно P(X < 6.2)?
- **Зад. 6.**: Средната стойност на теглото на новородените е 3500g със стандартно отклонение 500g. Каква е вероятността новороденото да тежи по-малко от 3100g?
- **Зад. 7.** Две мишки са тренирани да намерят изход в лабиринт. Първата това го прави със очаквана скорост X = N(80, 10^2). Втората за Y = N(78, 13^2). Каква е вероятността A да бъде по-бърза?