

Generación de números aleatorios

Modelos probabilísticos

Pruebas de bondad de ajuste

Kevin Steven García
Cesar Andres Saavedra

Universidad del Valle
Estadística
Simulación Estadística

Abril 2018

Introducción

Distribucion Poisson:

La distribución de Poisson es, una de las mas importantes distribuciones de probabilidad discretas, es decir, aquellas de las cuales pueden tomar los valores $0,1,2,3,\dots,k$.

Cada una de las Variables aleatorias representa el numero total de ocurrencias de un fenómeno durante un periodo de tiempo fijo o en una región fija del espacio y expresa la probabilidad de un numero k de ocurrencias.

Distribucion Logistica:

La distribución Logística es una distribución de probabilidad continua cuya función de distribución es la función logística, que aparece en el contexto de la regresión logística. Siendo la regresión logística un instrumento estadístico de análisis multivariado, de uso explicativo y predictivo. Que resulta útil cuando se tiene una variable dependiente dicotomica, atributo de ausencia o presencia que se puntuá con valores de 1 y 0, respectivamente.

Artículos

- Evidence in support of seismic hazard following Poisson distribution.

Pruebas de apoyo al riesgo sísmico tras la distribución Poisson.

- APLICAÇÃO DE REGRESSÃO LOGÍSTICA E ALGORITMOS GENÉTICOS NA ANÁLISE DE RISCO DE CRÉDITO.

Aplicación de la regresión logística y de los algoritmos genéticos en el análisis del riesgo de crédito.

Simulación Distribuciones

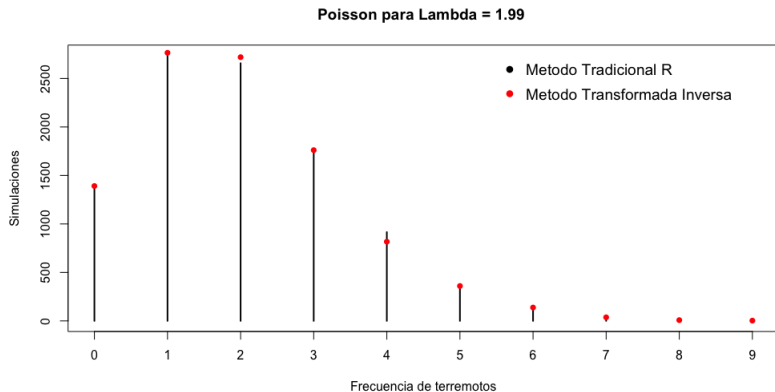


Figura: Comparación simulaciones en R y con el método propio de la distribución Poisson($\lambda = 1,99$)

Simulación Distribuciones

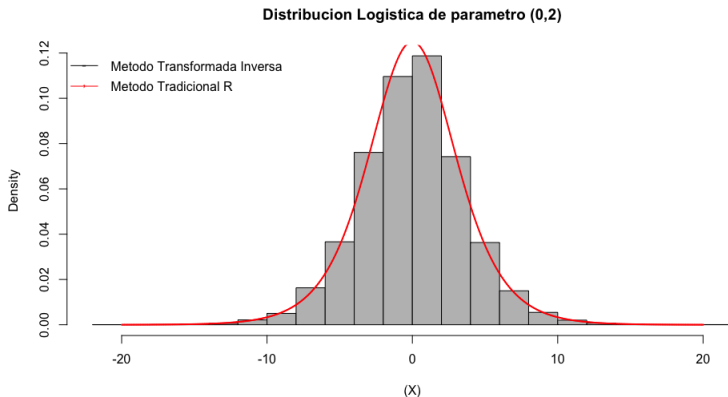


Figura: Comparación simulaciones en R y con el método de la transformada inversa para la distribución Logistica($a = 0, b = 2$)

Simulación Pruebas de bondad de ajuste

OBSERVACIÓN	OBS	ESP	$(OBS - ESP)^2 / ESP$
0	0.1404	0.1367	0.000100
1	0.2733	0.2720	0.000006
2	0.2617	0.2706	0.000292
3	0.1761	0.1795	0.000064
4	0.0933	0.0893	0.000179
5	0.0390	0.0355	0.000345
6	0.0125	0.0118	0.000041
7	0.0033	0.0033	0
8	0.0002	0.0008	0.00045
9	0.0002	0.0001	0.0001

El estadístico de prueba de la prueba χ^2 de Pearson nos dio

$$\chi_{obs}^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i} = 0,0015792$$

El p - valor = $P(\chi_{k-m-1}^2 > \chi_{obs}^2) = P(\chi_8^2 > 0,0015792) = 1.$

Simulación Pruebas de bondad de ajuste

```
>  
>  
>  
> #Prueba de bondad de ajuste kolmogorov para la logistic:  
> ks.test(X,'plogis',0,2)
```

One-sample Kolmogorov-Smirnov test

```
data: X  
D = 0.0040832, p-value = 0.9737  
alternative hypothesis: two-sided
```

```
>  
>  
>
```

Figura: Resultados prueba de Kolmogorov para los valores generados con el método de la transformada inversa para la Logistic($a = 0$, $b = 2$)

Conclusiones

- Por un lado, la distribución Poisson se utilizó para tratar de modelar y mitigar el riesgo sísmico dado que estos no son predecibles pero con ayuda de la probabilidad es posible medir el grado de incertidumbre al que se ven expuestas las personas, las estructuras y las grandes ciudades, así reducir el daño en zonas específicas donde el riesgo es mucho mayor.
- Por otro lado, la distribución Logistic se usó para predecir u obtener la probabilidad por medio de un modelo logit, de que un cliente sea buen o mal pagador en cuanto a sus créditos personales dadas ciertas características, esto es muy importante en las empresas financieras, ya que con tan solo unas pocas preguntas personales al cliente, pueden obtener dicha probabilidad logrando disminuir considerablemente los riesgos crediticios, evitando así generar inmensas perdidas en la empresa.

Conclusiones

- Con respecto a los métodos de simulación y pruebas de bondad de ajuste utilizadas, podemos ver que tanto el método utilizado para la distribución Poisson (método propio) como para la Logística (método de la transformada inversa) son muy acertados como generadores de variables aleatorias, ya que generan casi idénticamente datos que siguen dichas distribuciones, y esto se ve apoyado con las pruebas de bondad de ajuste utilizadas, ya que ambas se nos rechazan por mucho, es decir, los datos generados muestran demasiada evidencia de que siguen las distribuciones hipotéticas.

Bibliografía

Statistical distributions, Catherine Forbes, Merran Evans, Nicholas Hastings, Brian Peacock, Fourth edition.

<http://www.redalyc.org/html/750/75040605/>

<https://www.sciencedirect.com.bd.univalle.edu.co>

Anexos

Generadores de variables aleatorias

Generador variables aleatorias Poisson:

```
N=nsim
X<-c()
for (j in 1:N) {
  U<-runif(1)
  i=0 ; F=P=exp(-1)
  while (U>=F) {
    P=(1/(i+1))*P
    F=F+P
    i=i+1
  }
  X[j]=i
}
```

Anexos

Generadores de variables aleatorias

Generador de variables aleatorias Logistic:

$a=0$

$b=2$

$U \leftarrow \text{runif}(\text{nsim2}, 0, 1)$

$X = a + b \cdot \log((U)/(1-U))$