EYP1113 - Probabilidad y Estadística

Laboratorio 8

Profesoras: Natalia Venegas y Pilar Tello

Facultad de Matemáticas
Departamento de Estadística
Pontificia Universidad Católica de Chile

Segundo Semestre 2019

Contenido I

1 QQplot Weibull

2 Teorema del límite central

3 Estadísticos de orden

QQplot Weibull

Usando el siguiente código, se puede obtener las estimaciones para β (shape) y η (scale) de la distribución Weibull.

```
QQ.Weibull = function(y){
x=sort(y)
N=length(y)
p=(1:N)/(N+1)
plot(log(x)~log(-log(1-p)),
pch = 20, col = "darkblue", bty = "n", las = 1,
main = expression("QQ-Weibull"),
ylab = expression(log(x[p])),
xlab = expression(log(-log(1-p))))
abline(lm(log(x) ~ log(-log(1-p))), lwd = 3, col = "darkorange")
aux = lm(log(x) ~ log(-log(1-p)))
aux
}

eta = as.numeric(exp(QQ.Weibull(X)$coef[1]))
beta = as.numeric(1/QQ.Weibull(X)$coef[2])
```

La función de probabilidad acumulada de la distribución Weibull es $1 - e^{-(x/\eta)^{\beta}}$ por lo que al igualar ésta expresión a p e intentar obtener una función lineal, se tiene

$$\begin{aligned} p &= 1 - e^{-\left(\frac{x}{\eta}\right)^{\beta}} \\ -log(1-p) &= \left(\frac{x}{\eta}\right)^{\beta} \\ log(\eta) + \frac{1}{\beta} log(-log(1-p)) &= log(x) \end{aligned}$$

Luego, la regresión lineal $log(x) \sim log(-log(1-p))$ entrega pendiente $= 1/\beta$ y la intercepto $= log(\eta)$.
Por lo tanto, $\hat{\beta} = 1/\text{pendiente}$ y $\hat{\eta} = exp(\text{intercepto})$

Simule una muestra aleatoria Weibull y estime sus parámetros. Grafique su histograma y agregue linea de densidad Weibull con los parámetros obtenidos.

Teorema del límite central

El teorema dice que si $X_1, \ldots, X_n$ son variables aleatorias independientes e idénticamente distribuidas (iid) con

$$E(X_i) = \mu \quad \text{y} \quad \text{Var}(X_i) = \sigma^2$$

para todo $i = 1, \ldots, n$.

Entonces, cuando $n \to \infty$

$$Z_n = \frac{\displaystyle\sum_{i=1}^{n} X_i - n \cdot \mu}{\sqrt{n}\,\sigma} = \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \to Z \sim \text{Normal}(0, 1),$$

En otras palabras,

$$\sum_{i=1}^{n} X_i \overset{\text{aprox}}{\sim} \text{Normal}\left(n\,\mu,\ \sqrt{n}\,\sigma\right) \quad \text{o} \quad \frac{1}{n}\sum_{i=1}^{n} X_i \overset{\text{aprox}}{\sim} \text{Normal}\left(\mu,\ \frac{\sigma}{\sqrt{n}}\right)$$

Ejemplos

- Sean $X_1, \ldots, X_n \overset{\text{iid}}{\sim}$ Exponencial(λ), entonces

$$S_n = \sum_{i=1}^{n} X_i \sim \text{Gamma}(n,\ \lambda)$$
$$\overset{\text{aprox.}}{\sim} \text{Normal}\left(\frac{n}{\lambda},\ \frac{\sqrt{n}}{\lambda}\right)$$

- Sean $X_1, \ldots, X_n \overset{\text{iid}}{\sim}$ Poisson(λ), entonces

$$S_n = \sum_{i=1}^{n} X_i \sim \text{Poisson}(n\,\lambda)$$
$$\overset{\text{aprox.}}{\sim} \text{Normal}\left(n\,\lambda,\ \sqrt{n\,\lambda}\right)$$

Simule muestras aleatorias de los ejemplos anteriores (sumas). Grafique sus histogramas y agregue las líneas de densidad exacta y aproximada.

Estadísticos de orden

Sean $X_1, \ldots, X_n$ variables aleatorias continuas independientes e idéntimanete distribuidas con función densidad f_X y de distribución F_X, y sean $X_{(1)}, \ldots, X_{(n)}$ las mismas variables aleatorias, pero ordenadas de menor a mayor.
Estas variables aleatorias toman el nombre de estadísticos de orden y cada una es una función del vector aleatorio $(X_1, \ldots, X_n)'$
La función densidad del $k-$ésimo estadístico de orden está dada por

$$f_{X_{(k)}}(x) = \frac{n!}{(k-1)!(n-k)!}[F_X(x)]^{k-1}[1 - F_X(x)]^{n-k} f_X(x)$$

Estadísticos de orden

Para una muestra $X_1, \ldots, X_n$ de esta distribución, se definen:

$$Y_n = \max\{X_1, \ldots, X_n\}, \quad Y_1 = \min\{X_1, \ldots, X_n\}$$

La función de densidad del Y_n esta dada por:

$$f_{Y_n} = n\ [F_X(y)]^{n-1}\ f_X(y)$$

Mientras que la función de densidad de Y_1 corresponde a:

$$f_{Y_1} = n\ [1 - F_X(y)]^{n-1}\ f_X(y)$$

Por ejemplo, si $X_1, ..., X_n \sim Exponencial(\nu)$ iid, entonces

$$Y_{(1)} \sim Exponencial(n\nu)$$