Universidade Vigo

Departamento de Estatística e Investigación Operativa

E.Superior de Enxeñería Informática
Edificio Politécnico
Campus de Ourense
E-32004 Ourense

Tel. 986 387 000
http://esei.uvigo.es

Estatística

Apelidos:	**Nome:**	**DNI:**

1. (*3 puntos*) El archivo adjunto *galicia.txt*[1], contiene datos participación (personas censadas y frecuencia relativa de participación) de las elecciones municipales del 2011 y 2015 por ayuntamiento y provincia gallega. Para este conjunto de datos:

   a) Clasificar estadísticamente las variables del archivo.

   b) Obtén la diferencia de participación por ayuntamiento (expresando la variable en %). Dar la distribución completa de frecuencias agrupando la variable anterior con puntos de corte cada 2.5 % (desde -100 % hasta 100 %).

   c) Calcular la media muestral y la mediana con la variable agrupada y sin agrupar.

   d) Compara con un diagrama de cajas los valores por provincia. Extrae conclusiones.

2. *(2.5 puntos)* El nº de fallos de escritura con teclado (mecanografía) por página sigue una distribución de $Poisson(\lambda)$. Se clasifican los usuarios en 3 niveles *bajo, intermedio, alto* y se sabe que el número medio de fallos por página es 35, 20, 10, respectivamente. Responder

   a) ¿Cuál es la probabilidad de que un usuario de nivel intermedio cometa más de 110 errores en un documento de 6 páginas?

   b) Si un usuario comete 15 fallos, ¿a qué nivel se incluiría? Responder con criterios de probabilidad.

   c) Suponiendo que en un laboratorio el porcentaje de usuarios de nivel bajo, intermedio y alto es 35 %, 45 % y 20 %, respectivamente. ¿Cuál es la probabilidad de que un usuario elegido al azar cometa más de 15 fallos?

3. *(2 puntos)* Si $X \backsim N(0.5, 2)$, calcula,

   a) $P(-0.5 < X < 2) =$

   b) El valor de a que verifica $P(X > a) = 0.95$. Llega al valor de a usando las funciones de R relacionadas con la Normal con y sin los argumentos *mean=0, sd=1*.

4. *(2.5 puntos)* El lenguaje de programación *gfortran* dispone de una rutina *rand()* que devuelve números aleatorios supuestamente independientes entre 0 y 1, es decir, una v.a. continua con función de densidad:

$$f(x) = \begin{cases} 1 \text{ cuando } 0 < x < 1 \\ 0 \text{ en otro caso} \end{cases}$$

   Se define la siguiente v.a. $Y = -10 * X + 5$. Describe la v. aleatoria Y, para ello indica y justifica los valores de la variable Y y obtén las $P(Y \leq x)$, con $x = -10, -9, .., -1, 0, 1, ..., 9, 10$.

[1] Descargar desde la url *https://dl.dropboxusercontent.com/u/29008031/galicia.txt*

Solución Examen Estatística - Xuño 2015

25 de junio de 2015

```
> setwd("d:/2015-06 Scripts - EST-ESEI/")
```

```
> galicia <- read.table("galicia.txt",header=TRUE,sep="\t",dec=".")
```

```
> #a)
```

```
> names(galicia)

[1] "Provincia"    "Ayuntamiento" "censo2015"    "votos2015"
[5] "censo2011"    "votos2011"
```

```
> #"Provincia" --- cualitativa nominal
```

```
> #"Ayuntamiento" --- cualitativa nominal
```

```
> #"censo2015" --- cuantitativa discreta, valores enteros -- como son moitos se pode considerar continu
```

```
> #"votos2015" --- cuantitativa continua, valores no intervalo unidade
```

```
> #"censo2011" --- cuantitativa discreta, valores enteros -- como son moitos se pode considerar continu
```

```
> #"votos2011" --- cuantitativa continua, valores no intervalo unidade
```

```
> #b)
```

```
> galicia$dif <- (galicia$votos2015 - galicia$votos2011)*100
```

```
> galicia$dif.cut <- cut(galicia$dif,breaks=seq(-100,100,by=2.5))
```

```
> ni <- table(galicia$dif.cut)
```

```
>          # táboa de distribución de frecuencias completa
```

```
>          N <- sum(ni)
>          fi <- ni/N
>          m.c <- seq(-100,97.5,by=2.5) + 1.25
>          amp <- rep(2.5,80)
>          round(cbind( Marca.Clase=m.c, fi=fi, ni=ni, Ni=cumsum(ni), Fi= cumsum(fi), hi= fi/amp),2)

             Marca.Clase   fi  ni  Ni   Fi   hi
(-100,-97.5]      -98.75 0.00   0   0 0.00 0.00
(-97.5,-95]       -96.25 0.00   0   0 0.00 0.00
(-95,-92.5]       -93.75 0.00   0   0 0.00 0.00
(-92.5,-90]       -91.25 0.00   0   0 0.00 0.00
(-90,-87.5]       -88.75 0.00   0   0 0.00 0.00
(-87.5,-85]       -86.25 0.00   0   0 0.00 0.00
(-85,-82.5]       -83.75 0.00   0   0 0.00 0.00
(-82.5,-80]       -81.25 0.00   0   0 0.00 0.00
(-80,-77.5]       -78.75 0.00   0   0 0.00 0.00
(-77.5,-75]       -76.25 0.00   0   0 0.00 0.00
(-75,-72.5]       -73.75 0.00   0   0 0.00 0.00
(-72.5,-70]       -71.25 0.00   0   0 0.00 0.00
(-70,-67.5]       -68.75 0.00   0   0 0.00 0.00
(-67.5,-65]       -66.25 0.00   0   0 0.00 0.00
(-65,-62.5]       -63.75 0.00   0   0 0.00 0.00
(-62.5,-60]       -61.25 0.00   0   0 0.00 0.00
(-60,-57.5]       -58.75 0.00   0   0 0.00 0.00
(-57.5,-55]       -56.25 0.00   0   0 0.00 0.00
(-55,-52.5]       -53.75 0.00   0   0 0.00 0.00
(-52.5,-50]       -51.25 0.00   0   0 0.00 0.00
(-50,-47.5]       -48.75 0.00   0   0 0.00 0.00
(-47.5,-45]       -46.25 0.00   0   0 0.00 0.00
(-45,-42.5]       -43.75 0.00   0   0 0.00 0.00
(-42.5,-40]       -41.25 0.00   0   0 0.00 0.00
(-40,-37.5]       -38.75 0.00   0   0 0.00 0.00
(-37.5,-35]       -36.25 0.00   0   0 0.00 0.00
(-35,-32.5]       -33.75 0.00   0   0 0.00 0.00
(-32.5,-30]       -31.25 0.00   0   0 0.00 0.00
(-30,-27.5]       -28.75 0.00   0   0 0.00 0.00
(-27.5,-25]       -26.25 0.00   0   0 0.00 0.00
(-25,-22.5]       -23.75 0.00   0   0 0.00 0.00
(-22.5,-20]       -21.25 0.00   0   0 0.00 0.00
(-20,-17.5]       -18.75 0.00   0   0 0.00 0.00
(-17.5,-15]       -16.25 0.00   0   0 0.00 0.00
(-15,-12.5]       -13.75 0.01   3   3 0.01 0.00
(-12.5,-10]       -11.25 0.02   6   9 0.03 0.01
(-10,-7.5]         -8.75 0.09  28  37 0.12 0.04
(-7.5,-5]          -6.25 0.25  78 115 0.37 0.10
(-5,-2.5]          -3.75 0.34 108 223 0.71 0.14
(-2.5,0]           -1.25 0.19  61 284 0.90 0.08
(0,2.5]             1.25 0.07  23 307 0.98 0.03
```

```
(2.5,5]              3.75 0.01   4 311 0.99 0.01
(5,7.5]              6.25 0.00   1 312 0.99 0.00
(7.5,10]             8.75 0.00   1 313 1.00 0.00
(10,12.5]           11.25 0.00   0 313 1.00 0.00
(12.5,15]           13.75 0.00   0 313 1.00 0.00
(15,17.5]           16.25 0.00   0 313 1.00 0.00
(17.5,20]           18.75 0.00   0 313 1.00 0.00
(20,22.5]           21.25 0.00   0 313 1.00 0.00
(22.5,25]           23.75 0.00   0 313 1.00 0.00
(25,27.5]           26.25 0.00   0 313 1.00 0.00
(27.5,30]           28.75 0.00   0 313 1.00 0.00
(30,32.5]           31.25 0.00   0 313 1.00 0.00
(32.5,35]           33.75 0.00   0 313 1.00 0.00
(35,37.5]           36.25 0.00   0 313 1.00 0.00
(37.5,40]           38.75 0.00   0 313 1.00 0.00
(40,42.5]           41.25 0.00   0 313 1.00 0.00
(42.5,45]           43.75 0.00   0 313 1.00 0.00
(45,47.5]           46.25 0.00   0 313 1.00 0.00
(47.5,50]           48.75 0.00   0 313 1.00 0.00
(50,52.5]           51.25 0.00   0 313 1.00 0.00
(52.5,55]           53.75 0.00   0 313 1.00 0.00
(55,57.5]           56.25 0.00   0 313 1.00 0.00
(57.5,60]           58.75 0.00   0 313 1.00 0.00
(60,62.5]           61.25 0.00   0 313 1.00 0.00
(62.5,65]           63.75 0.00   0 313 1.00 0.00
(65,67.5]           66.25 0.00   0 313 1.00 0.00
(67.5,70]           68.75 0.00   0 313 1.00 0.00
(70,72.5]           71.25 0.00   0 313 1.00 0.00
(72.5,75]           73.75 0.00   1 314 1.00 0.00
(75,77.5]           76.25 0.00   0 314 1.00 0.00
(77.5,80]           78.75 0.00   0 314 1.00 0.00
(80,82.5]           81.25 0.00   0 314 1.00 0.00
(82.5,85]           83.75 0.00   0 314 1.00 0.00
(85,87.5]           86.25 0.00   0 314 1.00 0.00
(87.5,90]           88.75 0.00   0 314 1.00 0.00
(90,92.5]           91.25 0.00   0 314 1.00 0.00
(92.5,95]           93.75 0.00   0 314 1.00 0.00
(95,97.5]           96.25 0.00   0 314 1.00 0.00
(97.5,100]          98.75 0.00   0 314 1.00 0.00
```

```
> #c)
> mean(galicia$dif)

[1] -3.805446

> sum(ni*m.c)/N

[1] -3.789809
```

```
> median(galicia$dif)

[1] -4.064999
```

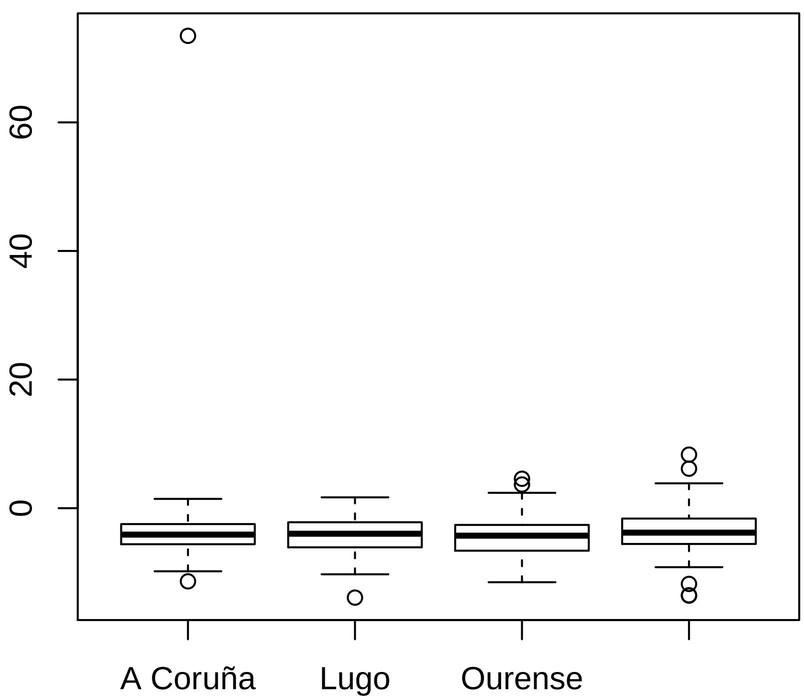
```
>                     which.min( cumsum(fi)<=0.5); cumsum(fi)[c(38,39)]

(-5,-2.5]
       39
(-7.5,-5] (-5,-2.5]
0.3662420 0.7101911

> m.c[38]+1.25 +(0.5-cumsum(fi)[38])/fi[39]* amp[39]

(-7.5,-5]
-4.027778
```

```
> #d)
>  boxplot(galicia$dif~ galicia$Provincia)
```



Universida de Vigo

Departamento de Estatística e Investigación Operativa

E.Superior de Enxeñería Informática
Edificio Politécnico
Campus de Ourense
E-32004 Ourense

Tel. 986 387 000
http://esei.uvigo.es

- **Solución:** Sea X_b, X_i, X_a = *nº de fallos según el nivel* $\sim Pois(\lambda_i)$. Como $E(X_i) = \lambda_i$, tenemos que $\lambda_b = 35, \lambda_i = 20, \lambda_a = 10$

  La variable X_i^5=*Nº de fallos en 6 páginas* $\sim Pois(6 * \lambda_i)$ dado que podemos suponer independencia en el nº de errores por página. Entonces:

  $$P(X_i^5 > 100) = 1 - ppois(110, 6 * 20) = 0.8061032$$

  □

  Si $X_? = 15$, entonces las probabilidades son:

  $$P(X_b = 15) = dpois(15, lambda = 35) = 6.985768e - 05$$
  $$P(X_i = 15) = dpois(15, lambda = 20) = 0.05164885$$
  $$P(X_a = 15) = dpois(15, lambda = 10) = 0.03471807$$

  Por lo tanto, la catalogación más probable es X_i.

  $$P(X > 15) = P(X_b > 15) * P(b) + P(X_i > 15) * P(i) + P(X_a > 15) * P(a) = (1 - ppois(15, lambda = 35)) * 0.35 + (1 - ppois(15, lambda = 20)) * 0.45 + (1 - ppois(15, lambda = 10)) * 0.20 = 0.7392757$$

-

$P(-0.5 < X < 2) =$

Solución: $P\left(-0.5 < X < 2\right) = P\left(X < 2\right) - P\left(X < -0.5\right) = P\left(N(0,1) < \frac{2-0.5}{2}\right) - P\left(N(0,1) < \frac{-0.5-0.5}{2}\right) = P\left(N(0,1) < 0.75\right)$ $P\left(N(0,1) < -0.5\right) = (1 - 0.2266) - P\left(N(0,1) > 0.5\right) = 0.7734 - 0.3085 = 0.4649.$ □

El valor de a que verifica $P(X > a) = 0.95$. Llega al valor de a usando las funciones de R relacionadas con la Normal con y sin los argumentos *mean=0, sd=1*.

Solución:

El cuantil 0.05, $x_{0.05}$, verifica que: $P\left(X < x_{0.05}\right) = 0.05$, y aplicando el complementario tenemos el valor de a. Tipificando y resolviendo obtenemos que:
$P\left(N(0,1) < \frac{x_{0.05}-0.50}{2}\right) = 0.05$. El cuantil de la normal 0.5 es negativo porque su probabilidad asociada es menor que 0.5. Aplicando simetría con respecto al cero y buscando en las tablas el valor 0.05 obtenemos que:
$-\frac{x_{0.05}-0.50}{2} = 1.645$, resolviendo, $x_{0.05} = 0.50 - 2 * 1.645 = -2.79$. □

La transformación de Y es lineal y si $x \in (0,1)$, entonces, $-10x \in (-10,0)$. Y por lo tanto $-10 * x + 5 \in (-5,5)$.
Las probabilidades se obtienen invirtiendo la inecuación:

$$
\begin{aligned}
P(Y \leq -10) &= 0 \\
P(Y \leq i) &= 0 \text{con } i = -9, -8, -7, -6, -5 \\
P(Y \leq -4) = P(X \in [0.9,1)) &= punif(1, min = 0, max = 1) - punif(0.9, min = 0, max = 1) = 0.1 \\
P(Y \leq -3) = P(X \in [0.8,1)) &= punif(1, min = 0, max = 1) - punif(0.8, min = 0, max = 1) = 0.2 \\
P(Y \leq -2) = P(X \in [0.7,1)) &= punif(1, min = 0, max = 1) - punif(0.7, min = 0, max = 1) = 0.3 \\
P(Y \leq -1) = P(X \in [0.6,1)) &= punif(1, min = 0, max = 1) - punif(0.6, min = 0, max = 1) = 0.4 \\
P(Y \leq 0) = P(X \in [0.5,1)) &= punif(1, min = 0, max = 1) - punif(0.5, min = 0, max = 1) = 0.5 \\
P(Y \leq 1) = P(X \in [0.4,1)) &= punif(1, min = 0, max = 1) - punif(0.4 min = 0, max = 1) = 0.6 \\
P(Y \leq 2) = P(X \in [0.3,1)) &= punif(1, min = 0, max = 1) - punif(0.3, min = 0, max = 1) = 0.7 \\
P(Y \leq 3) = P(X \in [0.2,1)) &= punif(1, min = 0, max = 1) - punif(0.2, min = 0, max = 1) = 0.8 \\
P(Y \leq 4) = P(X \in [0.1,1)) &= punif(1, min = 0, max = 1) - punif(0.1, min = 0, max = 1) = 0.9 \\
P(Y \leq 5) = P(X \in [0.0,1)) &= punif(1, min = 0, max = 1) - punif(0.0 min = 0, max = 1) = 1.0 \\
P(Y \leq i) &= 1 \text{con } i = 9, 8, 7, 6
\end{aligned}
$$