Estadísticos de orden

Hasta el momento nos hemos concentrado en estadísticos asociados con la localización $(\bar{x})$ y la dispersión (s^2) de los datos, pero no son los únicos que resultan útiles. Veremos otros asociados con la posición dentro de la muestra.

Def Sea $X_1, \dots, X_n$ una m.a. con f.d.p. $f_X(x;\theta)$ y f.p.a. $F_X(x;\theta)$.
Sin pérdida de generalidad, podemos ordenar la muestra de tal forma que

$$X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$$

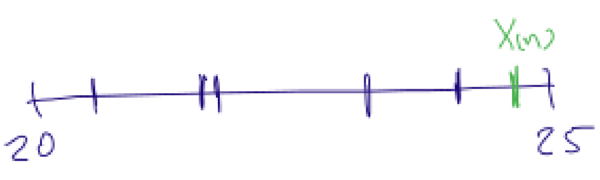
Al estadístico $T(\underline{X}) = X_{(i)}$ se le llama *i-ésimo estadístico de orden*

¿De qué nos sirven? $\longrightarrow$ p. ej. mediana
cuartiles
rango $X_{(n)} - X_{(1)}$
rango intercuartílico

¿Cómo se distribuyen?

Caso 1: máximo de la muestra

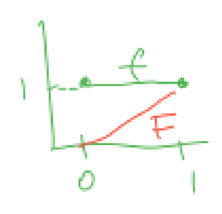
Queremos encontrar la distribución (muestral) del estadístico
$T(\underline{X}) = \text{máx}\{X_1, \dots, X_n\} = X_{(n)}$



$X_{(n)} < x$

$$
\begin{aligned}
F_{X_{(n)}}(x) &= P(X_{(n)} \leq x) \\
&= P(\text{máx}\{X_1, \dots, X_n\} \leq x) \\
&= P(X_1 \leq x, \dots, X_n \leq x) \\
&\overset{ind}{=} P(X_1 \leq x) \cdots P(X_n \leq x) \\
&= \prod_{i=1}^{n} P(X_i \leq x) \\
&= \prod_{i=1}^{n} F_X(x) = \left[F_X(x)\right]^n
\end{aligned}
$$

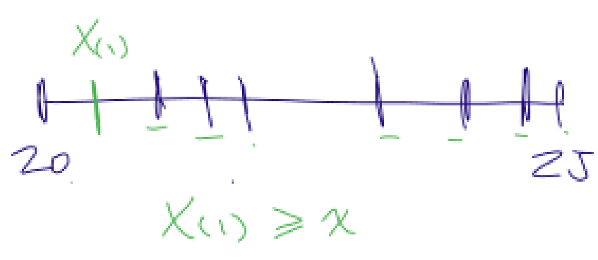
$X \sim U(0,1)$
$f_X(x) = 1$
$F_X(x) = x$



$F_{X_{(n)}}(x) = x^n$
$f_{X_{(n)}}(x) = n x^{n-1} \cdot 1$

$$f_{X_{(n)}}(x) = \frac{d}{dx} F_{X_{(n)}}(x) = n\left[F_X(x)\right]^{n-1} \cdot f_X(x)$$

Caso 2: mínimo de la muestra



$X_{(1)} \geq x$

$$
\begin{aligned}
F_{X_{(1)}}(x) &= P(X_{(1)} \leq x) \\
&= 1 - P(X_{(1)} > x) \\
&= 1 - P(\text{mín}\{X_1, \dots, X_n\} > x) \\
&= 1 - P(X_1 > x, \dots, X_n > x) \\
&\overset{ind}{=} 1 - P(X_1 > x) \cdots P(X_n > x) \\
&= 1 - \prod_{i=1}^{n} P(X_i > x) \\
&= 1 - \prod_{i=1}^{n} \left[1 - P(X_i \leq x)\right] \\
&= 1 - \prod_{i=1}^{n} \left[1 - F_X(x)\right] \\
&= 1 - \left[1 - F_X(x)\right]^n
\end{aligned}
$$

$f_{X_{(n)}}(x) = n(1-x)^{n-1}$

$$
\begin{aligned}
\Rightarrow f_{X_{(n)}}(x) &= \frac{d}{dx} F_{X_{(n)}}(x) = \ominus n\left[1 - F_X(x)\right]^{n-1} \cdot (-f_X(x)) \\
&= n\left[1 - F_X(x)\right]^{n-1} \cdot f_X(x)
\end{aligned}
$$

[Caso 3 (que no veremos) encontrar la distribución de $X_{(i)}$ para $1 \neq i \neq n$]

Ejemplo.

En una sucursal bancaria se desea estimar el tiempo mínimo que tarda un ejecutivo de cuenta en atender a clientes que llegan con el objetivo de abrir una nueva cuenta. Se sabe que la probabilidad de que al menos un cliente llegue para abrir una nueva cuenta en un día dado es del 75%. También se sabe que el número de clientes que llegan a la sucursal a abrir una cuenta en un día dado sigue un modelo Poisson con parámetro λ.

Sea X el tiempo que tarda un ejecutivo en abrir una cuenta nueva para un cliente.

(a) ¿Qué modelo de probabilidad es adecuado para X?
(b) Con la información provista, calcular el valor de λ.
(c) Calcular la f.d.p. de $X_{(1)}$ para una muestra $X_1, \dots, X_n$.
(d) ¿Cuál es la probabilidad de que el tiempo mínimo que deba esperar un cliente para concluir su trámite no sobrepase los 20 minutos para una muestra de tamaño 30?

$$[\lambda] = \frac{\text{personas}}{\text{tiempo}}$$

$$\left[\frac{1}{\lambda}\right] = \frac{\text{tiempo}}{\text{persona}}$$

$\xrightarrow{a}$ $X \sim Exp\left(\frac{1}{\lambda}\right)$

$\xrightarrow{b}$ $W \sim Po(\lambda)$

W = # de clientes que llegan para abrir una cuenta (en un día)

$P(W \geq 1) = 0.75$

$P(W \leq x)$
$P(W = x)$

$$
\begin{aligned}
&\Rightarrow 1 - P(W < 1) = 0.75 \\
&\Rightarrow 1 - P(W = 0) = 0.75 \\
&\Rightarrow 1 - \frac{e^{-\lambda} \cdot \lambda^0}{0!} = 0.75 \\
&\Rightarrow 0.25 = e^{-\lambda} \Rightarrow \ln(0.25) = -\lambda \\
&\Rightarrow \lambda = 1.38
\end{aligned}
$$

$\xrightarrow{c}$ Sea $Y = X_{(1)} = \text{mín}\{X_1, \dots, X_n\}$ donde $X_1, \dots, X_n \overset{iid}{\sim} Exp\left(\frac{1}{\lambda}\right)$

$\lambda = 1.38 \Rightarrow \frac{1}{\lambda} = 0.72$

$F_{X_i}(x) = P(X_i \leq x) = 1 - e^{-\frac{1}{\lambda}x}$

$P(X_i > x) = 1 - F_X(x) = e^{-\frac{1}{\lambda}x}$

$$
\begin{aligned}
F_Y(x) = F_{X_{(1)}}(x) &= 1 - \left[1 - P(X_i \leq x)\right]^{n-1} \\
&= 1 - \left[P(X_i > x)\right]^{n-1} \\
&= 1 - \left[e^{-\frac{1}{\lambda}x}\right]^n = 1 - e^{-0.72nx}
\end{aligned}
$$

$$f_Y(x) = f_{X_{(1)}}(x) = \frac{d}{dx} F_{X_{(1)}}(x) = -(-0.72n)e^{-0.72nx} = 0.72n e^{-0.72nx}$$

Alternativamente:
$$
\begin{aligned}
& n\left[1 - F_X(x)\right]^{n-1} f_X(x) \\
&= n\left[e^{-\frac{1}{\lambda}x}\right]^{n-1} \cdot \frac{1}{\lambda} \cdot e^{-\frac{1}{\lambda}x} \\
&= n\left[e^{-0.72x}\right]^{n-1} 0.72 e^{-0.72x} \\
&= 0.72n\left[e^{-0.72nx + \cancel{0.72x} - \cancel{0.72x}}\right] \\
&= 0.72n e^{-0.72nx}
\end{aligned}
$$

$\xrightarrow{d}$ 20 mín $\approx$ 0.014 días

$$
\begin{aligned}
& P(\text{mín}\{X_1, \dots, X_{30}\} \leq 20 \text{ mín}) \\
&= P(Y \leq 20 \text{ mín}) \\
&= P(Y \leq 0.014 \text{ días}) \\
&= F_Y(0.014) \\
&= 1 - e^{-0.72(30)(0.014)} = 0.26
\end{aligned}
$$

Distribuciones de muestreo para v.a. discretas y "pequeñas"

Ejemplo

La hembra de cierta especie de mamífero entra en periodo de celo una sola vez al año (puedo o no quedar preñada).
Sea $X = \#$ de crías que puede tener una madre de esta especie.
Suponer que X es tal que

x	0	1	2
$P(X=x)$	0.25	0.6	0.15

Se toma una muestra de tamaño 2 con reemplazo)
Obtener la distribución muestral de $\bar{X}$ y s^2.

(X_1, X_2)	Prob (conjunta)	$\bar{X}$	$s^2 = \frac{1}{n-1}\boxed{\sum (X_i - \bar{X})^2}$
(0, 0)	$(0.25)(0.25) = 0.0625$	$\frac{0+0}{2} = 0$	$(0-0)^2 + (0-0)^2 = 0$
(0, 1)	$(0.25)\cdot(0.6) = 0.15$	$\frac{0+1}{2} = 0.5$	$0.5 = (0-0.5)^2 + (1-0.5)^2 = 0.5^2 + 0.5^2 = 0.25 + 0.25$
(0, 2)	$(0.25)(0.15) = 0.0375$	1	2
(1, 0)	$(0.6)(0.25) = 0.15$	0.5	0.5
(1, 1)	$(0.6)(0.6) = 0.36$	1	0
(1, 2)	$(0.6)(0.15) = 0.09$	1.5	0.5
(2, 0)	$(0.15)(0.25) = 0.0375$	1	2
(2, 1)	$(0.15)(0.6) = 0.09$	1.5	0.5
(2, 2)	$(0.15)(0.15) = 0.0225$	2	0
	Total 1		

X_1+X_2	$\bar{x}$	$P(\bar{X}=x)$
0	0	0.0625
1	0.5	0.3
2	1	0.435
3	1.5	0.18
4	2	0.0225
	Total	1

s^2	$P(s^2 = x)$
0	0.445
0.5	0.48
2	0.075
Total	1

$n \longrightarrow \infty$

$\bar{X} \xrightarrow[n\to\infty]{} E[X]$

¿ $s^2 \xrightarrow[n\to\infty]{} V(X)$?