



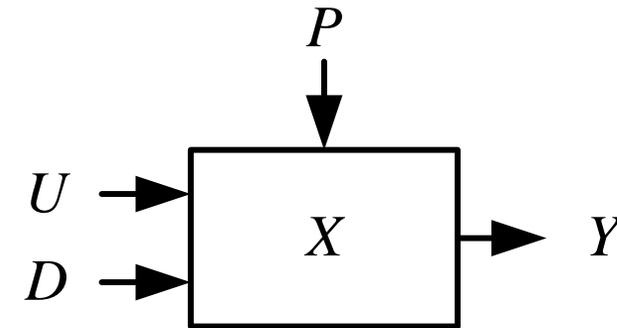
# Modelado de entradas Parte I

Enrique E. Tarifa, Facultad de Ingeniería, UNJu

# Modelado de entradas

# Clasificación de variables

- Parámetros ( $P$ )
- Variables de entrada:
  - Manipulables ( $U$ )
  - Perturbación ( $D$ )
- Variables de salida ( $Y$ )
- Variables internas ( $I$ )
- Variables de estado ( $X \subseteq I$ )



- Tiempo entre arribos de clientes
- Tipo de operación
- Monto de la operación

# Etapas del modelado de entradas

1. Colección de datos
2. Identificación de la distribución
3. Determinación de parámetros
4. Evaluación

# Resultados

Colección

- Tabla  $x$

Identificación

- $f(x)$  o  $p(x)$

Ajuste

- Parámetros

Evaluación

- Distribución

# Resultados

Colección

- Tabla  $x$

Identificación

- $f(x)$  o  $p(x)$

Ajuste

- Parámetros

Evaluación

- Distribución

# Colección de datos

- Mediciones, conocimiento de expertos, registros históricos del sistema en estudio o de sistemas similares.
- Tiempos entre eventos, duraciones, cantidades.
- $n$  depende del nivel de confianza de la evaluación.

$i$	$x$
1	$x_1$
2	$x_2$
...	...
$n$	$x_n$

# Resultados

Colección

- Tabla  $x$

Identificación

- $f(x)$  o  $p(x)$

Ajuste

- Parámetros

Evaluación

- Distribución

# Identificación de la distribución

- Determinar  $n$ ,  $x_{\min}$  y  $x_{\max}$
- Rango:  $r = x_{\max} - x_{\min}$
- Clases:  $c = n^{1/2}$
- Ancho:  $b = r/c$
- Clase  $k = 1, 2, \dots, c$
- Frecuencias absolutas  $f_k$
- Frecuencias relativa  $fr_k = f_k/n$
- Frecuencia acumulada  $fa_k$

# Histograma

$k$				
1				
2				
...				
$c$				

# Histograma



$x_{\min}$     $x_{\min}+b$     $x_{\min}+2b$     $x_{\max}$

$k$	$LS$			
1	$x_{\min}+b$			
2	$x_{\min}+2b$			
...	...			
$c$	$x_{\max}$			

Para evitar problemas por redondeo, usar  $x_{\max}$  al final.

# Histograma

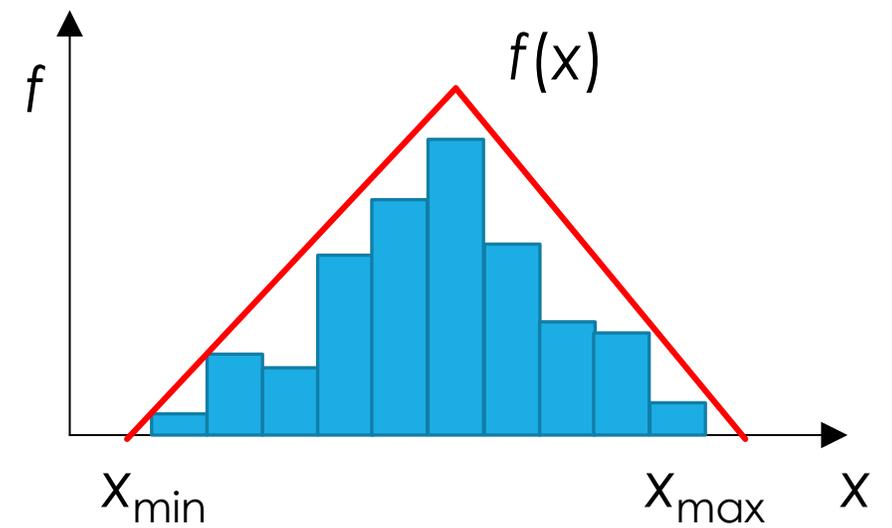
$k$	$LS$	$f$		
1	$x_{\min} + b$	$f_1$		
2	$x_{\min} + 2b$	$f_2$		
...	...	...		
$C$	$x_{\max}$	$f_C$		
<b>Total</b>		<b><math>n</math></b>		

# Histograma

$k$	$LS$	$f$	$fr$	
1	$x_{\min} + b$	$f_1$	$f_1/n$	
2	$x_{\min} + 2b$	$f_2$	$f_2/n$	
...	...	...	...	
$C$	$x_{\max}$	$f_c$	$f_c/n$	
<b>Total</b>		<b><math>n</math></b>	<b>1</b>	

# Histograma

$k$	$LS$	$f$	$fr$	$fa$
1	$x_{\min} + b$	$f_1$	$f_1/n$	$fr_1$
2	$x_{\min} + 2b$	$f_2$	$f_2/n$	$fa_1 + fr_2$
...	...	...	...	...
$C$	$x_{\max}$	$f_c$	$f_c/n$	1
<b>Total</b>		<b><math>n</math></b>	<b>1</b>	



# Mediciones

4.18	1.16	5.28	1.76	0.34	0.42	3.88	3.89	1.22	2.58
2.60	0.38	0.55	1.45	0.21	1.84	1.09	1.79	0.05	5.21
5.35	2.81	4.55	0.38	0.55	0.85	5.90	2.89	1.30	1.91
2.01	0.62	1.95	11.96	1.73	3.42	3.98	4.51	2.49	0.50
0.63	1.56	2.93	0.16	3.88	0.63	2.02	3.20	1.25	0.80
2.68	1.36	0.09	0.98	5.86	2.07	1.23	2.71	0.31	0.39
0.76	0.69	0.43	1.12	1.75	0.73	4.53	1.74	0.47	1.79
3.86	0.01	1.25	3.32	4.94	6.56	0.95	2.80	2.43	0.93
2.87	2.79	0.94	1.82	2.51	1.86	0.84	1.66	0.39	4.54
2.43	3.13	5.24	1.28	0.15	0.56	1.62	1.19	0.87	1.82

Ver archivo Distribución exponencial.xlsx

# Parámetros del histograma

- $x_{\min} = 0.01$
- $x_{\max} = 11.96$
- $r = 11.95$
- $n = 100$
- $c = 10$
- $b = 1.195$

4.18	1.16	5.28	1.76	0.34	0.42	3.88	3.89	1.22	2.58
2.60	0.38	0.55	1.45	0.21	1.84	1.09	1.79	0.05	5.21
5.35	2.81	4.55	0.38	0.55	0.85	5.90	2.89	1.30	1.91
2.01	0.62	1.95	11.96	1.73	3.42	3.98	4.51	2.49	0.50
0.63	1.56	2.93	0.16	3.88	0.63	2.02	3.20	1.25	0.80
2.68	1.36	0.09	0.98	5.86	2.07	1.23	2.71	0.31	0.39
0.76	0.69	0.43	1.12	1.75	0.73	4.53	1.74	0.47	1.79
3.86	0.01	1.25	3.32	4.94	6.56	0.95	2.80	2.43	0.93
2.87	2.79	0.94	1.82	2.51	1.86	0.84	1.66	0.39	4.54
2.43	3.13	5.24	1.28	0.15	0.56	1.62	1.19	0.87	1.82

# Tabla del histograma

- $x_{\min} = 0.01$
- $x_{\max} = 11.96$
- $r = 11.95$
- $n = 100$
- $c = 10$
- $b = 1.195$

<b><i>k</i></b>	<b><i>LS</i></b>	<b><i>f</i></b>	<b><i>fr</i></b>	<b><i>fa</i></b>
<b>1</b>	1.205	37	0.37	0.37
<b>2</b>	2.400	26	0.26	0.63
<b>3</b>	3.595	18	0.18	0.81
<b>4</b>	4.790	10	0.10	0.91
<b>5</b>	5.985	7	0.07	0.98
<b>6</b>	7.180	1	0.01	0.99
<b>7</b>	8.375	0	0.00	0.99
<b>8</b>	9.570	0	0.00	0.99
<b>9</b>	10.765	0	0.00	0.99
<b>10</b>	11.960	1	0.01	1.00
<b>Total</b>		<b>100</b>	<b>1.00</b>	

Archivo Inicio Insertar Disposición de página Fórmulas **Datos** Revisar Vista Programador Ayuda Comentarios Compartir

Obtener y transformar datos Actualizar todo Consultas y conexiones Consultas & conexiones  
 Cotizaciones Monedas Tipos de datos Ordenar Ordenar y filtrar Filtro Borrar Volver a aplicar Avanzadas Herramientas de datos Previsión Previsión Esquema

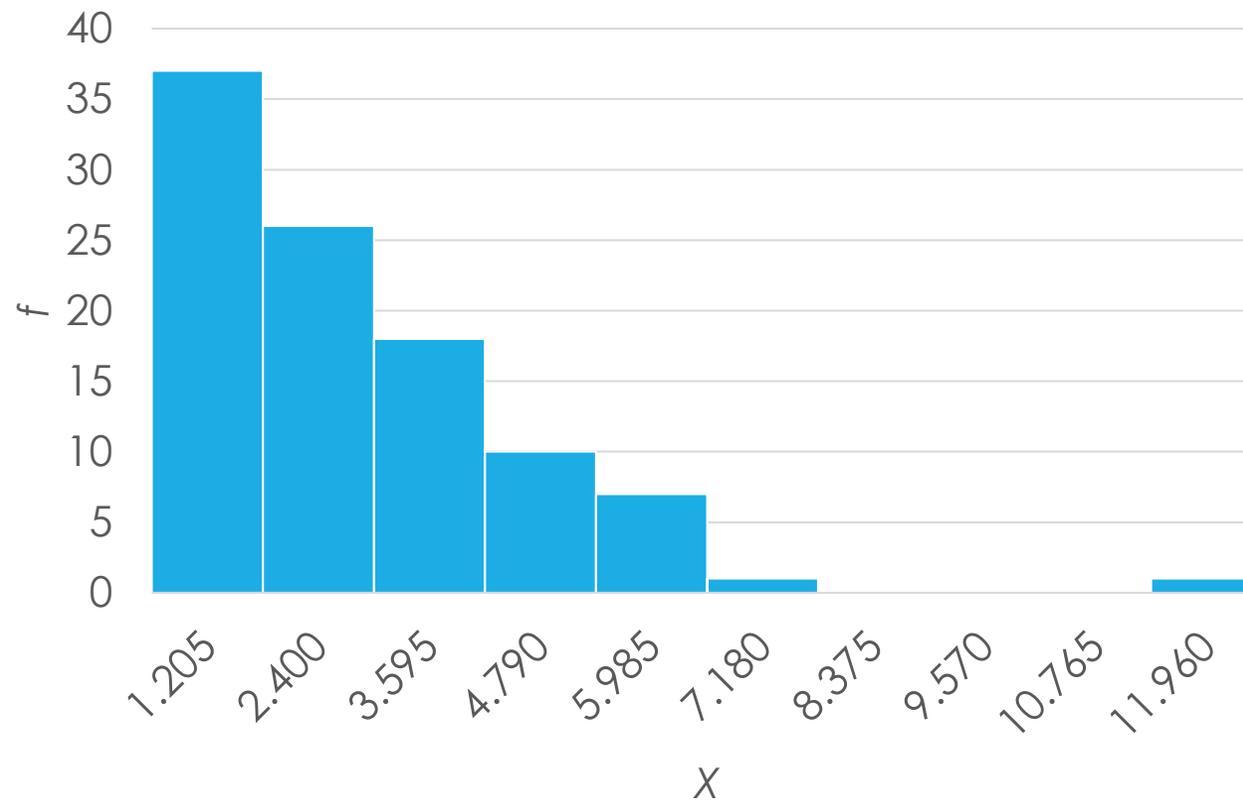
E8 fx

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	<i>i</i>	<i>X</i>		<b>Parámetros del histograma</b>			<i>k</i>	<i>LS</i>	<i>f</i>	<i>fr</i>	<i>fa</i>		Ctrl+Shift
2	1	4.18		n =	100		1	1.205	37	0.37	0.37		
3	2	2.60		c =	10		2	2.400	26	0.26	0.63		
4	3	5.35		Xmin =	0.01		3	3.595	18	0.18	0.81		
5	4	2.01		Xmax =	11.96		4	4.790	10	0.1	0.91		
6	5	0.63		r =	11.95		5	5.985	7	0.07	0.98		
7	6	2.68		b =	1.195		6	7.180	1	0.01	0.99		
8	7	0.76					7	8.375	0	0	0.99		
9	8	3.86		Xm =	2.14		8	9.570	0	0	0.99		
10	9	2.87		L =	0.4676175		9	10.765	0	0	0.99		
11	10	2.43					10	11.960	1	0.01	1		
12	11	1.16					<b>Total</b>		<b>100</b>	<b>1</b>			





# Histograma



# Frecuencias

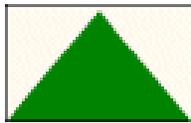
- El primer intervalo debe ser [LI,LS].
- Los siguientes deben ser (LI,LS].
- Fórmula matricial:
  - Seleccionar rango.
  - Editar primera celda.
  - Ctrl+Shift+Enter: Frecuencia absoluta
  - Ctrl+Enter: Frecuencia acumulada
- La nueva versión de Excel tiene fórmulas por desbordamiento.

# Distribuciones

$f(x)$  y  $p(x)$



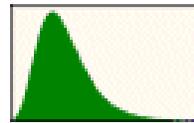
Normal



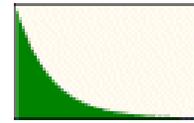
Triangular



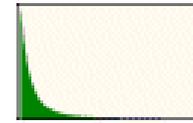
Uniform



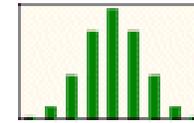
Lognormal



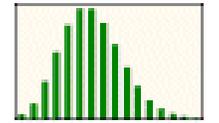
Exponential



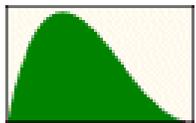
Pareto



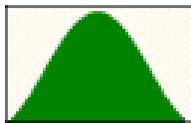
Binomial



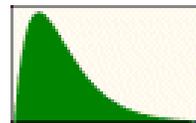
Poisson



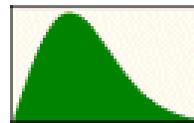
Beta



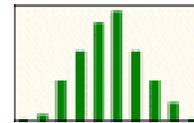
BetaPERT



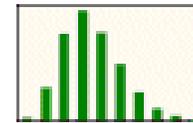
Gamma



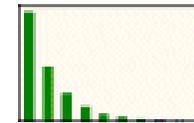
Weibull



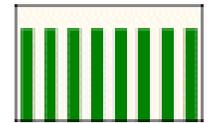
Hypergeometric



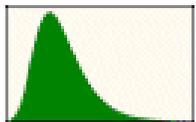
Neg Binomial



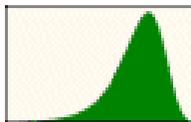
Geometric



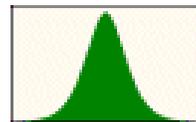
Discrete Uniform



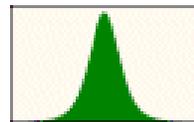
Max Extreme



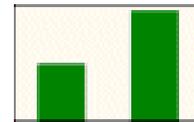
Min Extreme



Logistic



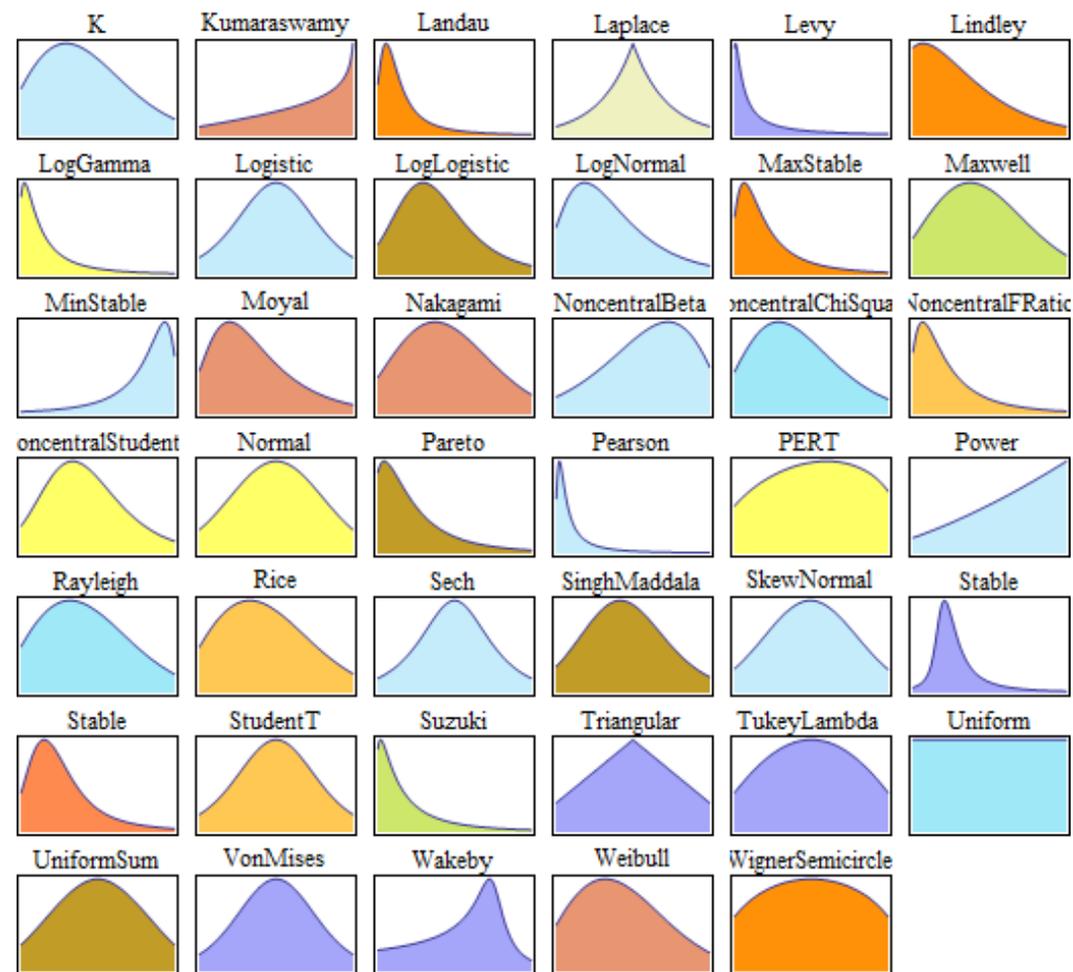
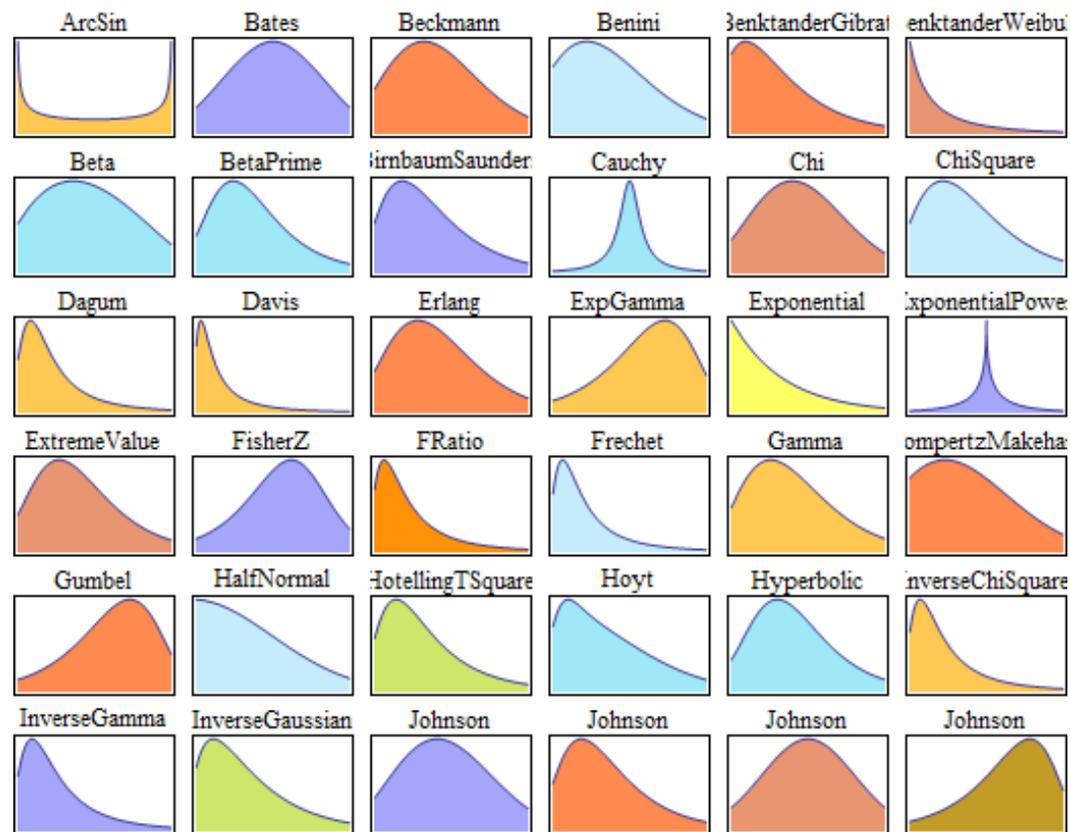
Student's t

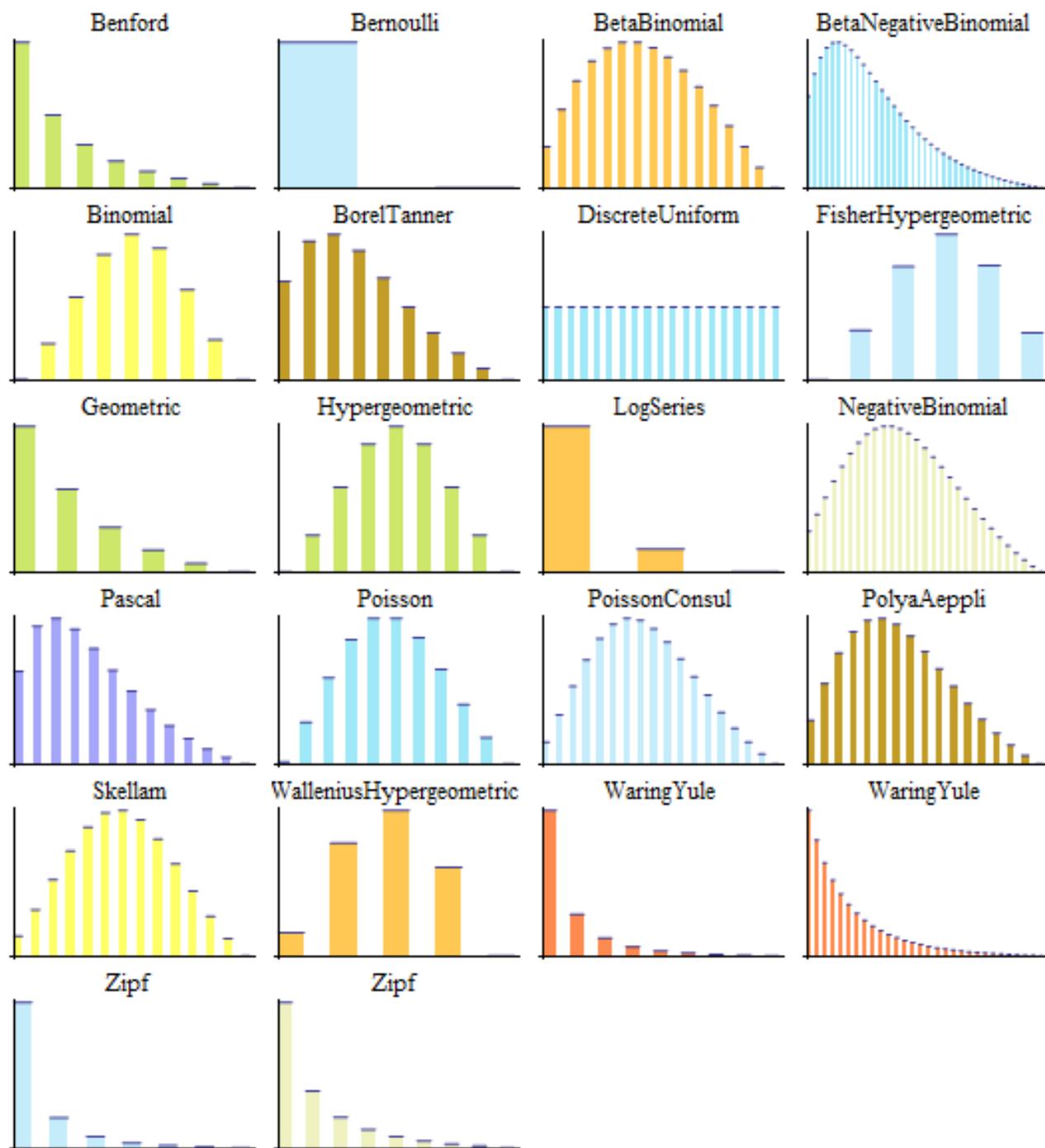


Yes-No

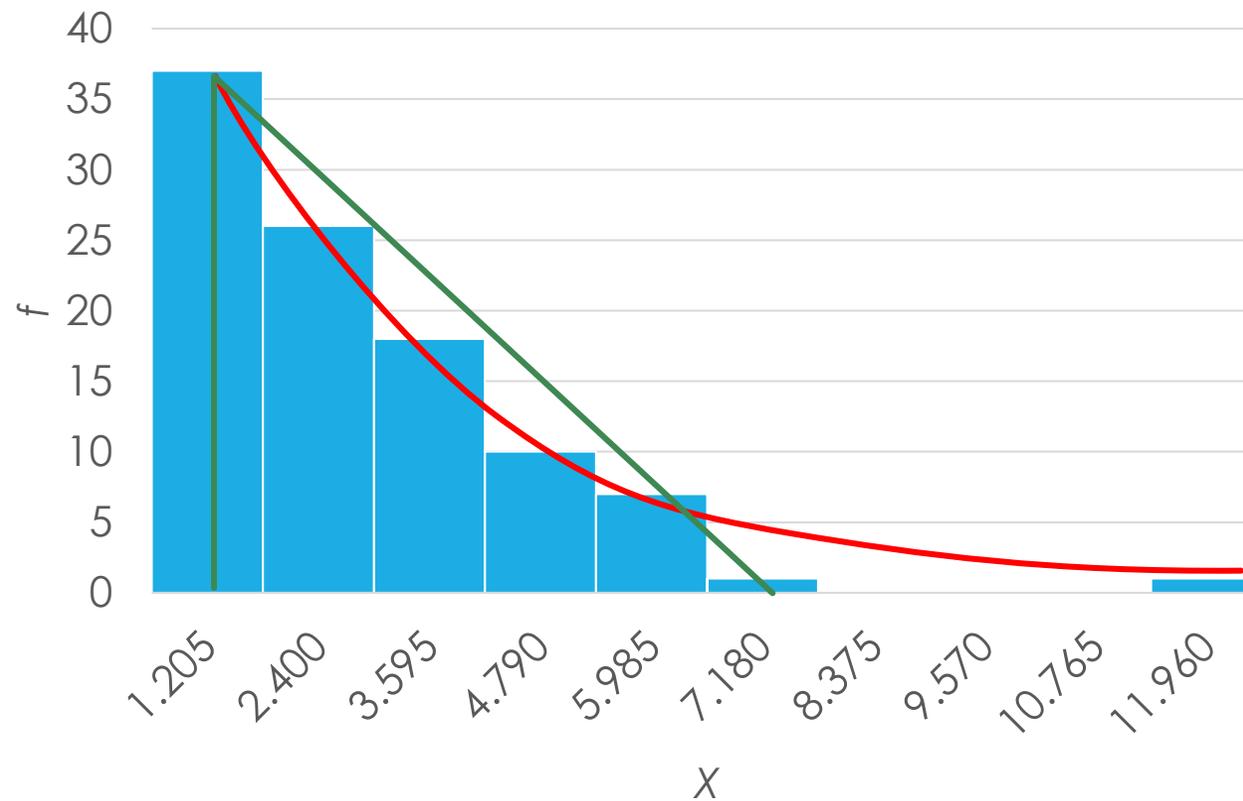


Custom





# Histograma



# Distribuciones

- Binomial. Modela el número de éxitos en  $n$  pruebas independientes con probabilidad de éxito  $p$ :
  - Número de pendrives defectuosos en un lote de  $n$ .
- Geométrica. Modela la cantidad de pruebas a realizar hasta encontrar el primer éxito en un lote con probabilidad de éxito  $p$ :
  - Cantidad de llamadas a realizar a un centro de ayuda hasta que atiendan.

# Distribuciones

- Poisson. Modela el número de eventos independientes que ocurren en una cantidad fija de tiempo o espacio:
  - Número de clientes que llegan a un centro comercial durante una hora, o el número de defectos encontrados en 30 m<sup>2</sup> de una lámina de metal.
- Normal. Modela la distribución de un proceso que puede representarse como la suma de varios procesos:
  - El tiempo de ensamblaje de un automóvil puede representarse como la suma de los tiempos de ensamblaje de las distintas partes.

# Distribuciones

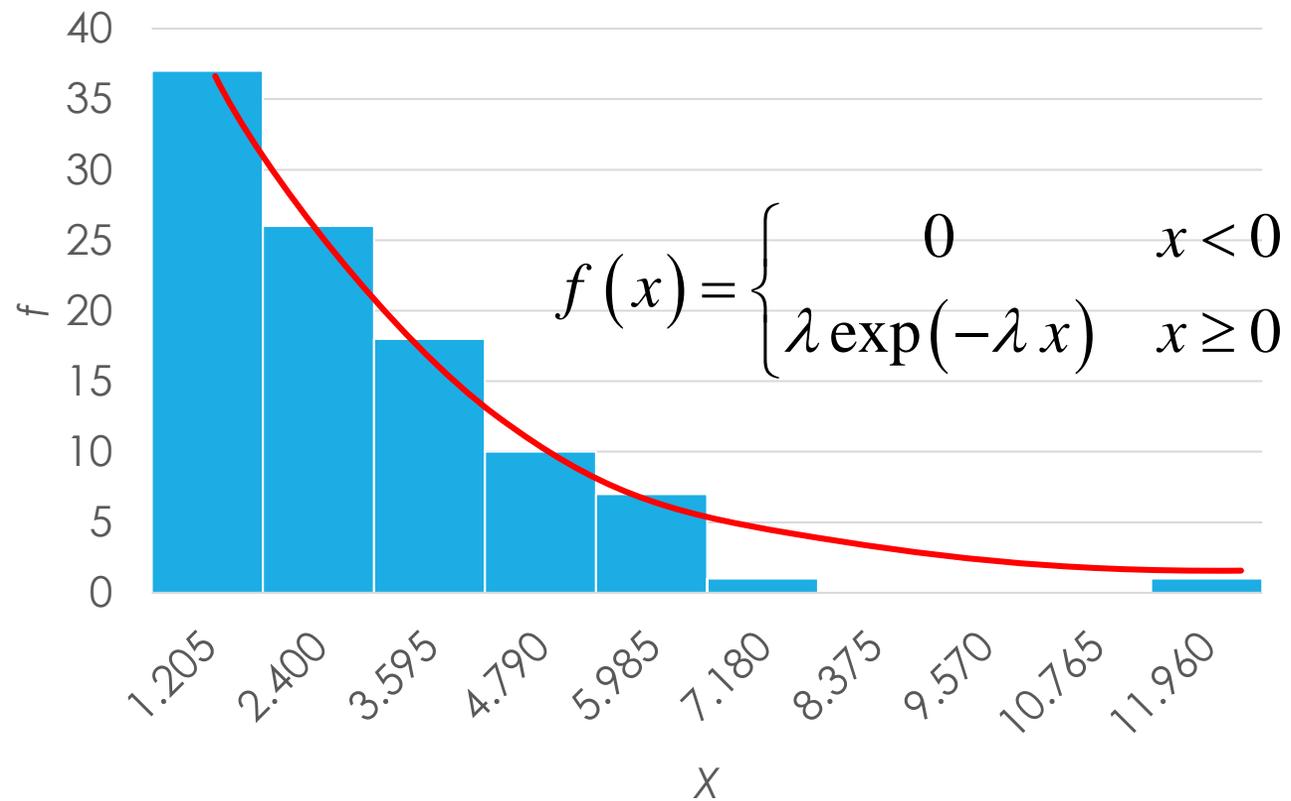
- Exponencial. Modela el tiempo entre eventos independientes, o tiempos de procesos sin memoria, donde no se puede inferir el tiempo final del proceso a partir del tiempo transcurrido:
  - Tiempos entre arribos de clientes.
- Erlang. Modela los tiempos de los procesos que pueden representarse como la suma de varios procesos con distribución exponencial:
  - Una red de computadora falla cuando una computadora y dos computadoras de respaldo fallan sucesivamente.

# Distribuciones

- Uniforme. Modela procesos completamente inciertos, ya que todos los valores de  $X$  son igualmente probables.
- Triangular: Modela procesos de los cuales solo se conocen los valores mínimo, máximo y más probable:
  - Se conoce la duración mínima, máxima y más probable de la calidad de un producto.

Si los datos son mediciones de tiempos de entre eventos, se prefiere a la distribución exponencial en lugar de la triangular.

# Histograma



# Mediciones

7	3	9	3	4	5	7	4	6	9
3	3	6	3	5	7	7	4	5	5
7	3	6	8	1	4	1	2	5	9
7	6	5	3	8	2	4	5	4	6
5	7	5	2	3	5	5	6	8	7
8	1	3	5	3	6	5	5	5	5
5	2	6	4	8	4	5	10	4	2
7	5	6	4	5	2	9	9	6	5
6	5	4	6	9	7	8	3	4	9
6	7	1	3	1	3	6	5	7	4

Ver archivo Distribución binomial.xlsx

# Parámetros del histograma

- $x_{\min} = 1$
- $x_{\max} = 10$
- $n = 100$
- $b = 1$

7	3	9	3	4	5	7	4	6	9
3	3	6	3	5	7	7	4	5	5
7	3	6	8	1	4	1	2	5	9
7	6	5	3	8	2	4	5	4	6
5	7	5	2	3	5	5	6	8	7
8	1	3	5	3	6	5	5	5	5
5	2	6	4	8	4	5	10	4	2
7	5	6	4	5	2	9	9	6	5
6	5	4	6	9	7	8	3	4	9
6	7	1	3	1	3	6	5	7	4

# Tabla del histograma

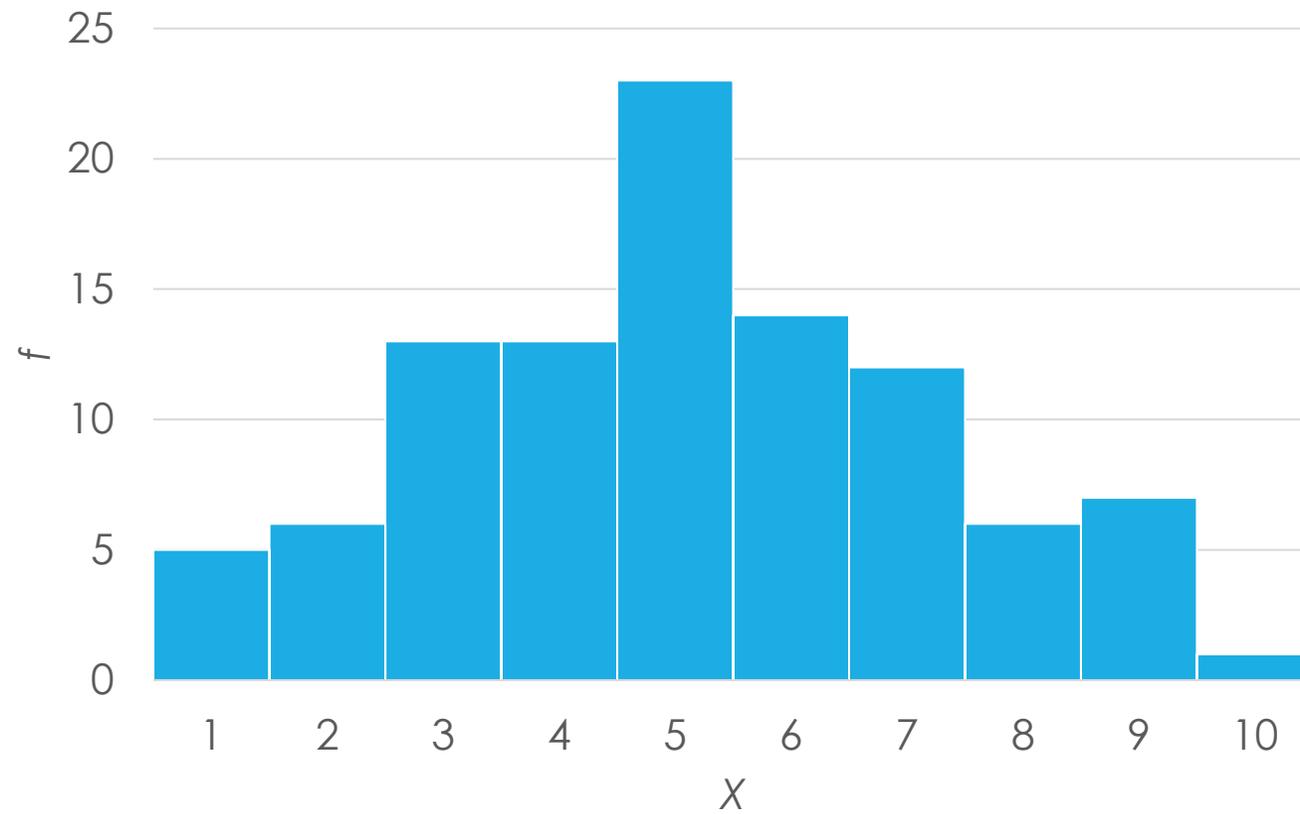
- $X_{\min} = 1$
- $X_{\max} = 10$
- $n = 100$
- $b = 1$

<i>k</i>	<i>LS</i>	<i>f</i>	<i>fr</i>	<i>fa</i>
1	1	5	0.05	0.05
2	2	6	0.06	0.11
3	3	13	0.13	0.24
4	4	13	0.13	0.37
5	5	23	0.23	0.6
6	6	14	0.14	0.74
7	7	12	0.12	0.86
8	8	6	0.06	0.92
9	9	7	0.07	0.99
10	10	1	0.01	1.00
<b>Total</b>		<b>100</b>	<b>1.00</b>	





# Histograma

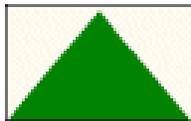


# Distribuciones

$$p(x)$$



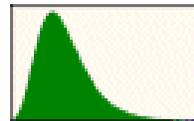
Normal



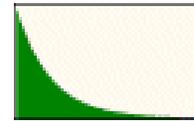
Triangular



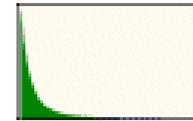
Uniform



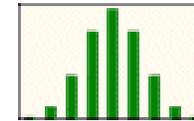
Lognormal



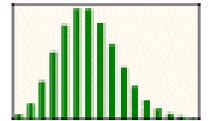
Exponential



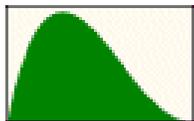
Pareto



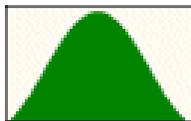
Binomial



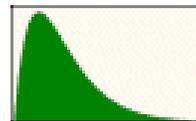
Poisson



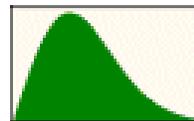
Beta



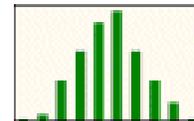
BetaPERT



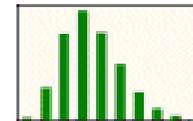
Gamma



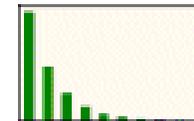
Weibull



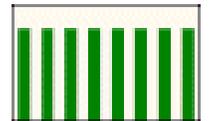
Hypergeometric



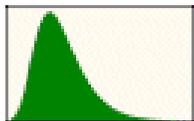
Neg Binomial



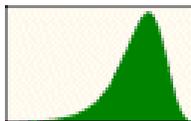
Geometric



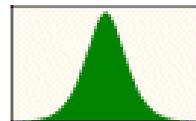
Discrete Uniform



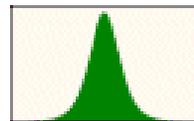
Max Extreme



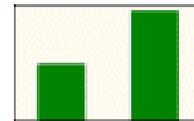
Min Extreme



Logistic



Student's t



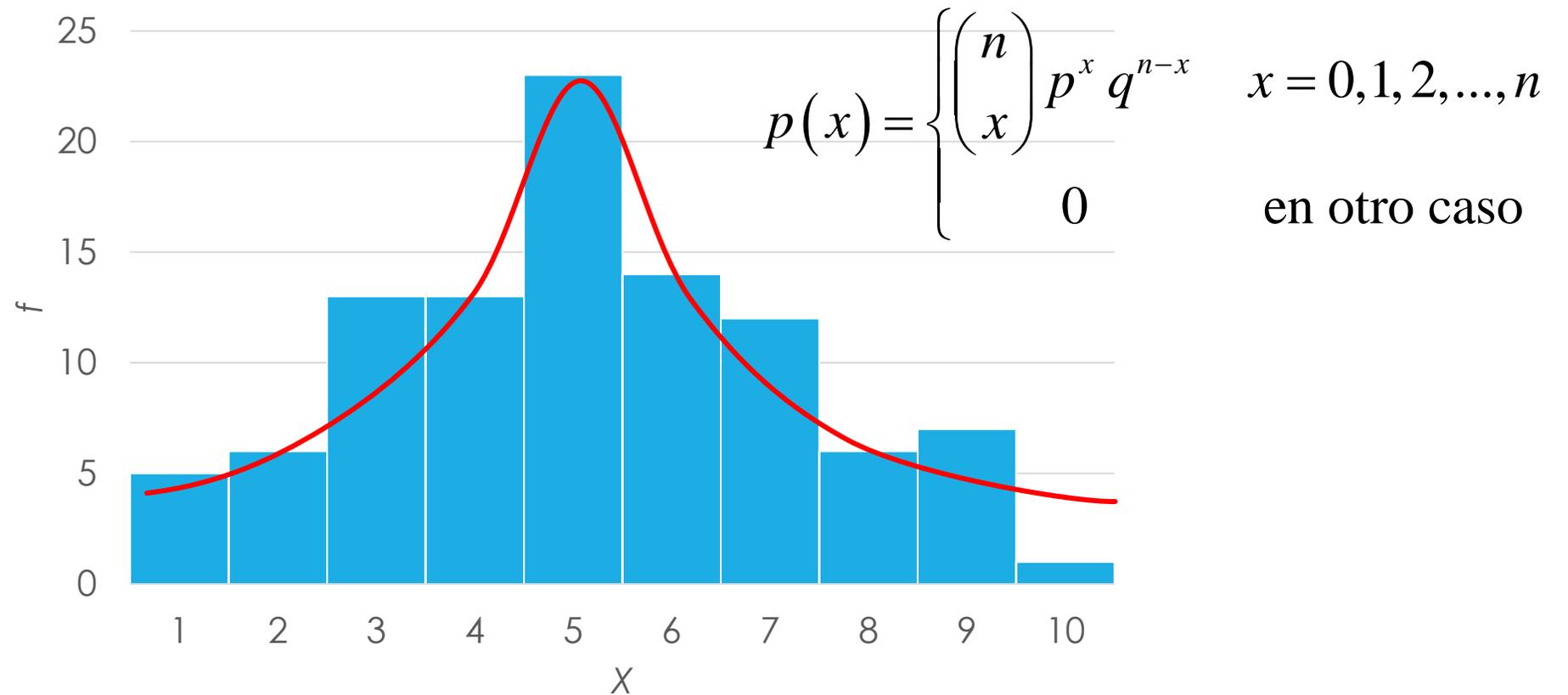
Yes-No



Custom

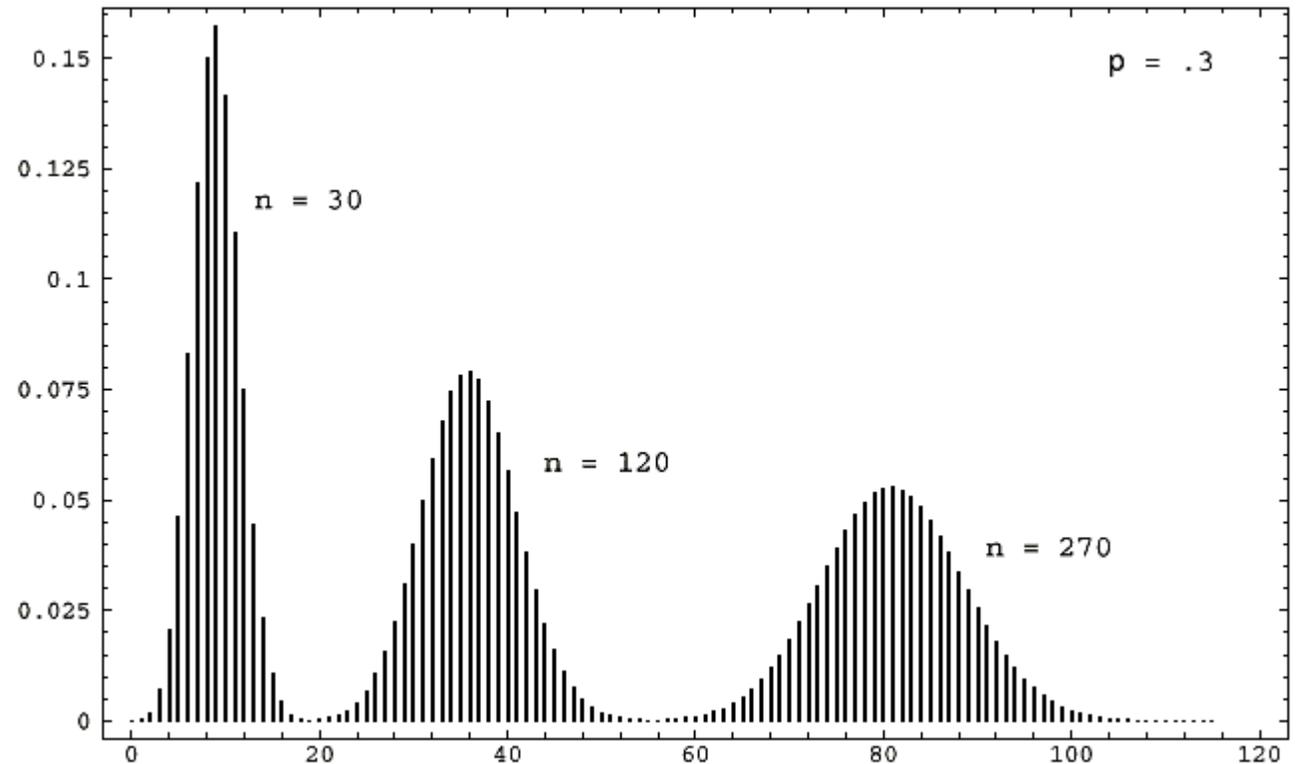
Cantidad de autos que no superan la RTV cuando se revisan 50 autos por día.

# Histograma



# Binomial( $n, p$ ) a Normal( $\mu, \sigma$ )

- Para una buena aproximación:
  - $n \geq 30$
  - $np \geq 5$
  - $nq \geq 5$
  - $p \approx q$
- Distribución normal:
  - $\mu = np$
  - $\sigma = \sqrt{npq}$



# Resultados

Colección

- Tabla  $x$

Identificación

- $f(x)$  o  $p(x)$

Ajuste

- Parámetros

Evaluación

- Distribución

# Estimación de parámetros

- Método de la máxima verosimilitud:

$$\text{Max}_P \sum_{i=1}^n \ln(p(x_i, P))$$

$$\text{Max}_P \sum_{i=1}^n \ln(f(x_i, P))$$

# Estimación de parámetros

Estimadores:

$$\bar{X} = \frac{\sum_{i=1}^n x_i}{n}$$

$$S^2 = \frac{\sum_{i=1}^n (x_i - \bar{X})^2}{n-1} = \frac{\sum_{i=1}^n x_i^2 - n\bar{X}^2}{n-1}$$

$$\lim_{n \rightarrow \infty} \bar{X} \rightarrow E(X)$$

$$\lim_{n \rightarrow \infty} S^2 \rightarrow V(X)$$

En Excel:

- Promedio(rango)
- Var.S(rango)

# Estimación de parámetros

- Poisson:  $\alpha \approx \bar{X}$
- Binomial:  $p \approx \frac{\bar{X}}{n}$
- Geométrica:  $p \approx 1/\bar{X}$
- Exponencial:  $\lambda \approx 1/\bar{X}$
- Uniforme  $[0,b]$ :  $b \approx X_{\max} (n + 1)/n$
- Normal:  $\mu \approx \bar{X}, \sigma^2 = S^2$

# Distribución exponencial

Exponencial:  $\lambda \approx 1/\bar{X}$

Xm =	2.14
L =	0.4676175

Xm =	=PROMEDIO(B2:B101)
L =	=1/E9

# Distribución binomial

Binomial:  $p \approx \frac{\bar{X}}{n}$

Xm =	5.12
p =	0.05

Xm =	=PROMEDIO(B2:B101)
p =	=E9/E2

# Resultados

Colección

- Tabla  $x$

Identificación

- $f(x)$  o  $p(x)$

Ajuste

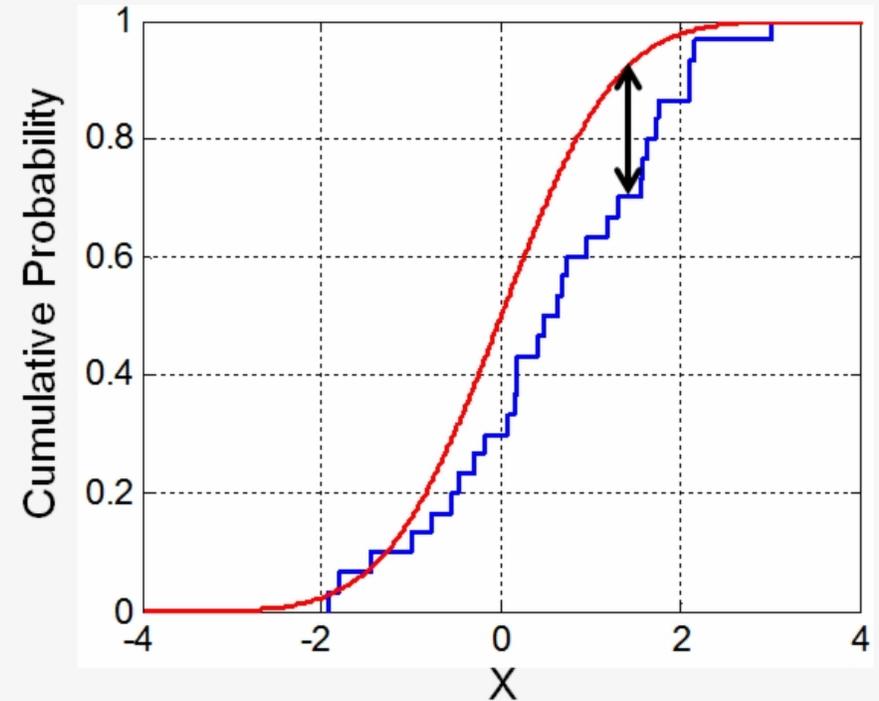
- Parámetros

Evaluación

- Distribución

# Evaluación del ajuste

- Kolmogorov-Smirnov: Decide en base a la máxima desviación entre la distribución acumulada teórica y la distribución acumulada empírica.



# Evaluación del ajuste

- Chi-cuadrado: Decide en base a la suma de errores al cuadrado que surgen de comparar el histograma con la distribución teórica.

$$\chi^2 = \sum_{k=1}^c \frac{(f_k - f_k^e)^2}{f_k^e}$$

# Evaluación del ajuste

- Si la muestra es pequeña, las pruebas aceptarán cualquier distribución.
- Si la muestra es grande, las pruebas rechazarán a todas las distribuciones propuestas.
- Por lo tanto, estas pruebas son solo un elemento más a tener en cuenta durante la evaluación.