Probabilidad (4)

Jesús García de Jalón de la Fuente

IES Ramiro de Maeztu
Madrid

2020

La distribución binomial o de Bernouilli

La distribución binomial es un modelo matemático del siguiente experimento:

- Una prueba tiene dos resultados posibles: éxito con probabilidad p y fracaso con probabilidad $q = 1 - p$.
- La prueba se repite n veces. Las pruebas son independientes.
- El resultado es el número de éxitos que se obtienen.

Sea $X =$ «número de éxitos obtenidos». X puede tomar los valores 0, 1, 2, …, n.

La función de probabilidad es:

$$\mathrm{p}(X = k) = \binom{n}{k} p^k q^{n-k}$$

El valor esperado es $\mathrm{E}(X) = np$ y la varianza $\mathrm{Var}(X) = npq$.

Si una variable aleatoria X tiene una función de probabilidad como esta, diremos que sigue la distribución binomial $\mathrm{B}(n, p)$ y escribiremos $X \sim \mathrm{B}(n, p)$.

Ejemplo

Supongamos que se lanza un dado 8 veces y contamos el número de ocasiones en las que se obtiene un múltiplo de 3.

Sea el éxito de la prueba obtener un múltiplo de 3 y X la variable que representa el número de éxitos.

$X \sim \mathrm{B}\left(8, \frac{1}{3}\right)$. Las probabilidades son las siguientes:

$$\mathrm{p}(X=0)=\binom{8}{0}\left(\frac{1}{3}\right)^0\left(\frac{2}{3}\right)^8=0{,}0390 \qquad \mathrm{p}(X=1)=\binom{8}{1}\left(\frac{1}{3}\right)^1\left(\frac{2}{3}\right)^7=0{,}1561 \qquad \mathrm{p}(X=2)=\binom{8}{2}\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^6=0{,}2731$$

$$\mathrm{p}(X=3)=\binom{8}{3}\left(\frac{1}{3}\right)^3\left(\frac{2}{3}\right)^5=0{,}2731 \qquad \mathrm{p}(X=4)=\binom{8}{4}\left(\frac{1}{3}\right)^4\left(\frac{2}{3}\right)^4=0{,}1707 \qquad \mathrm{p}(X=5)=\binom{8}{5}\left(\frac{1}{3}\right)^5\left(\frac{2}{3}\right)^3=0{,}0683$$

$$\mathrm{p}(X=6)=\binom{8}{6}\left(\frac{1}{3}\right)^6\left(\frac{2}{3}\right)^2=0{,}0171 \qquad \mathrm{p}(X=7)=\binom{8}{7}\left(\frac{1}{3}\right)^7\left(\frac{2}{3}\right)^1=0{,}0024 \qquad \mathrm{p}(X=8)=\binom{8}{8}\left(\frac{1}{3}\right)^8\left(\frac{2}{3}\right)^0=0{,}0002$$

El valor esperado es:

$$\mathrm{E}(X)=8\cdot\frac{1}{3}=\frac{8}{3}$$

La varianza y la desviación típica:

$$\mathrm{Var}(X)=8\cdot\frac{1}{3}\cdot\frac{2}{3}=\frac{16}{9}\,; \qquad \sigma=\sqrt{\mathrm{Var}(X)}=\frac{4}{3}$$

Diagrama de barras

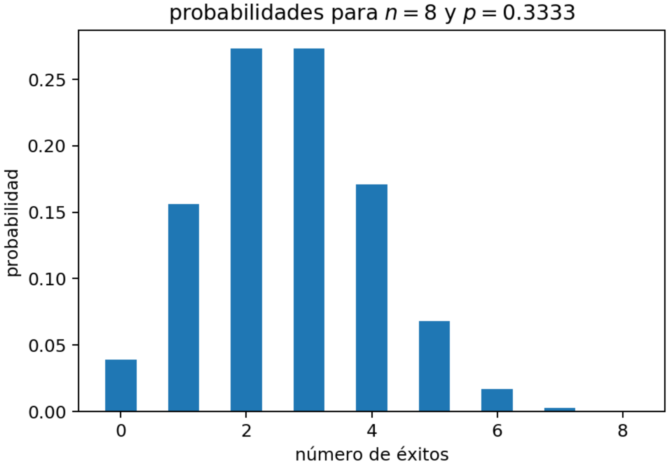
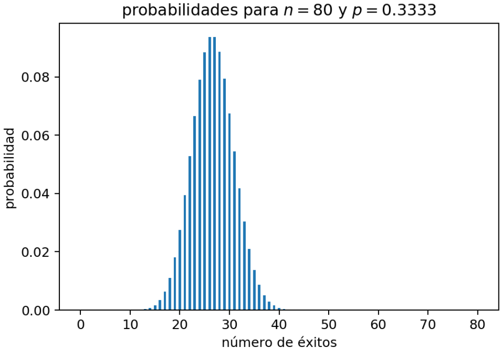


Diagrama de barras



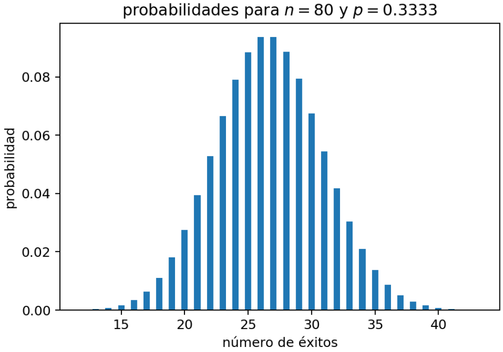
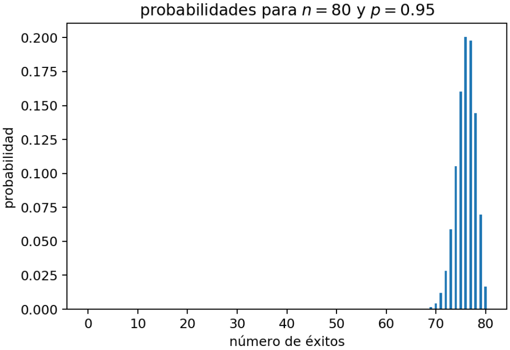
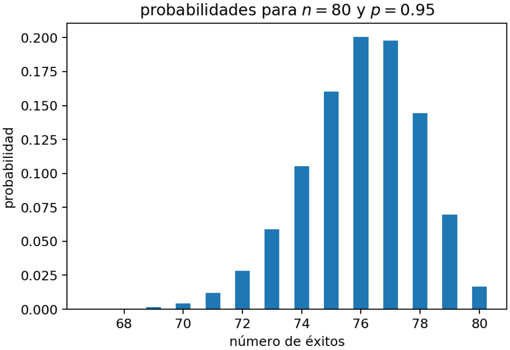


Diagrama de barras



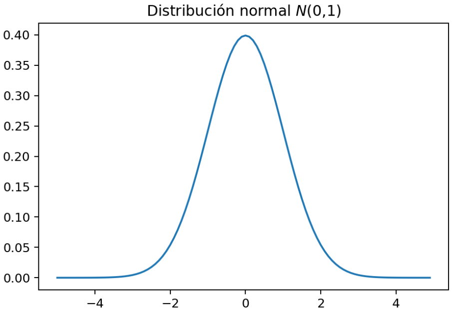


La distribución normal standard $\mathrm{N}(0,1)$ es una distribución de variable continua. La variable aleatoria se suele representar por Z y su función de densidad es;

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

La media o valor esperado de esta distribución es cero y la desviación típica es igual a 1.

La gráfica de esta función es:



Las probabilidades de la distribución normal son complicadas de obtener por integración.

Se suelen usar tablas con los valores de la función de distribución $\Phi(x)$.

En la tabla solamente aparecen los valores positivos de x. Para calcular los valores de $\Phi(x)$ para los x negativos hay que tener en cuenta que de la simetría de la función resulta que:

$$\Phi(-x) = 1 - \Phi(x)$$

Por ejemplo:

$$\begin{aligned} \mathrm{p}(1{,}25 \leq Z \leq 2{,}23) &= \Phi(2{,}23) - \Phi(1{,}25) \\ &= 0{,}9871 - 0{,}8944 \\ &= 0{,}0927 \end{aligned}$$

z	0.00	0.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

Haciendo el cambio $z = \mu + \sigma x$ se obtiene la distribución normal de media μ y desviación típica σ que se representa por $\mathrm{N}(\mu, \sigma)$. Su función de densidad es:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Las probabilidades de esta distribución se pueden obtener a partir de las de la distribución $\mathrm{N}(0, 1)$ haciendo el cambio:

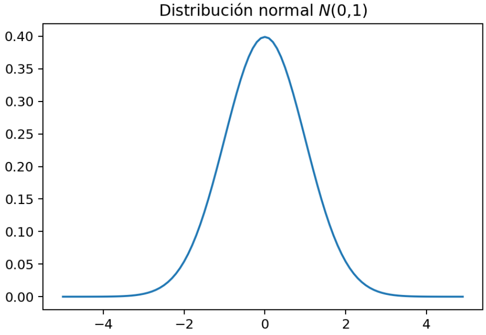
$$z = \frac{x - \mu}{\sigma}$$

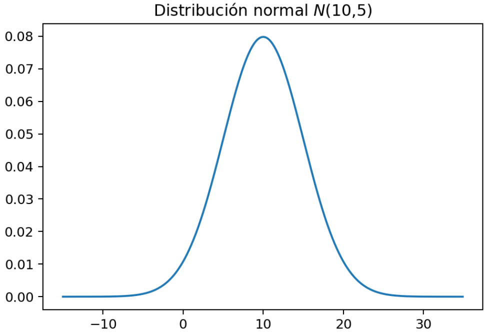
que se llama tipificar la variable.

Por ejemplo, sea $X \sim \mathrm{N}(50, 10)$:

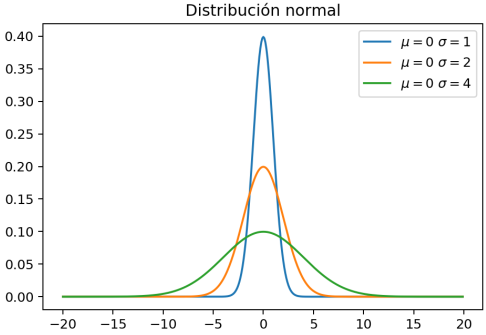
$$\mathrm{p}(X \leq 65) = \mathrm{p}\left(Z \leq \frac{65 - 50}{10}\right) = \mathrm{p}(Z \leq 1{,}5) = 0{,}9332$$

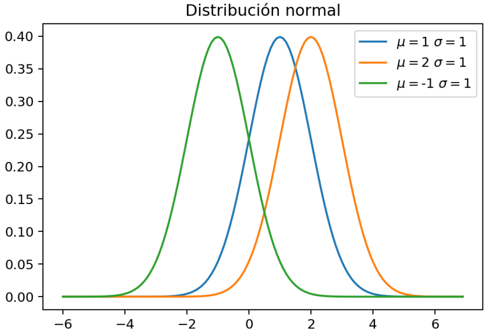
La distribución N(μ, σ)



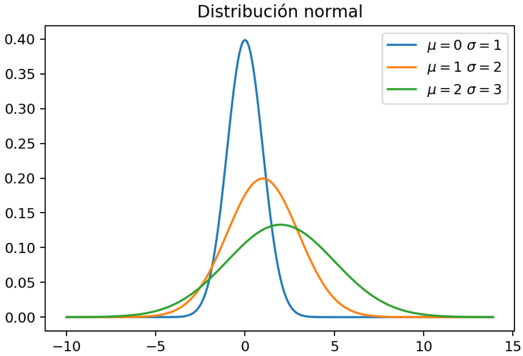


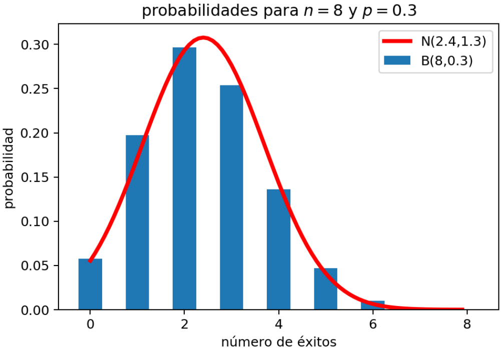
La distribución N(μ, σ)

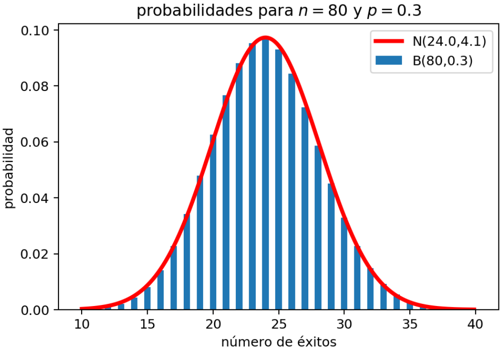


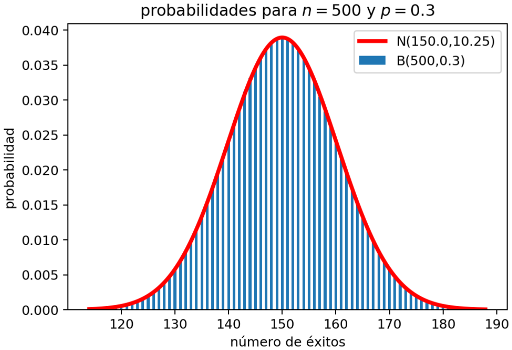


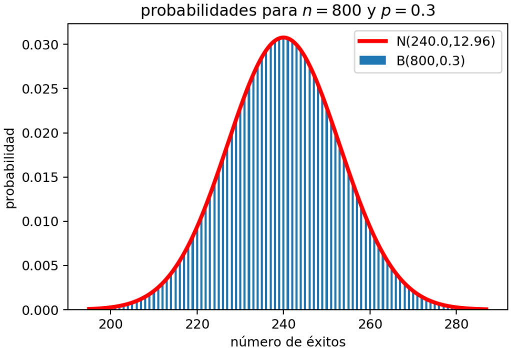
La distribución N(μ, σ)



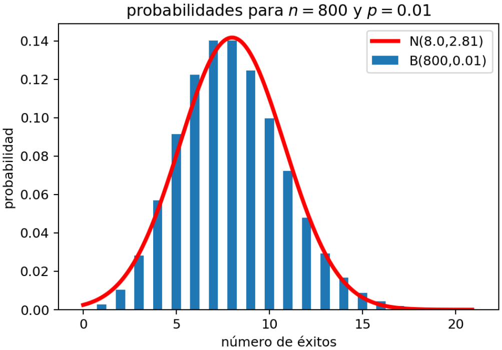

probabilidades para $n = 8$ y $p = 0.3$
N(2.4,1.3)
B(8,0.3)
probabilidad
número de éxitos
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0
2
4
6
8

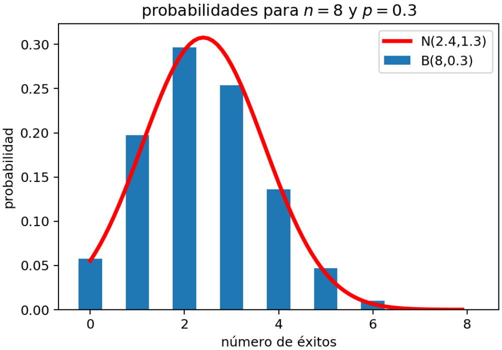

probabilidades para $n = 80$ y $p = 0.3$
N(24.0,4.1)
B(80,0.3)
probabilidad
número de éxitos
0.00
0.02
0.04
0.06
0.08
0.10
10
15
20
25
30
35
40


probabilidades para $n = 500$ y $p = 0.3$
N(150.0,10.25)
B(500,0.3)
probabilidad
número de éxitos
0.000
0.005
0.010
0.015
0.020
0.025
0.030
0.035
0.040
120
130
140
150
160
170
180
190


probabilidades para $n = 800$ y $p = 0.3$
N(240.0,12.96)
B(800,0.3)
probabilidad
número de éxitos
0.000
0.005
0.010
0.015
0.020
0.025
0.030
200
220
240
260
280

Distribuciones binomial y normal





Sea $X \sim \mathrm{B}(n, p)$ y $X' \sim \mathrm{N}(\mu, \sigma)$ con $\mu = np$ y $\sigma = \sqrt{npq}$. Sea $f(x)$ la función de densidad de $X' \sim \mathrm{N}(\mu, \sigma)$.

Si np y nq son grandes:

$$\mathrm{p}(X = k) \simeq f(k)$$

La aproximación se suele utilizar para calcular probabilidades acumuladas de la distribución binomial. En ese caso se aplica la corrección de continuidad o de Yates que asocia a cada número entero k el intervalo $\left[k - \frac{1}{2}, k + \frac{1}{2}\right]$ para calcular las probabilidades:

$$\mathrm{p}(X \leq k) \simeq \mathrm{p}\left(X' \leq k + \frac{1}{2}\right)$$
$$\mathrm{p}(k_1 \leq X \leq k_2) \simeq \mathrm{p}\left(k_1 - \frac{1}{2} \leq X' \leq k_2 + \frac{1}{2}\right)$$

Como regla práctica se suele decir que la aproximación es buena si np y nq son mayores que 5.

Agradecimiento

Gracias por vuestra atención