MATEMATICAS (MATS I y II)

Bachillerato

Distribuciones de probabilidad

Departamento de Matemáticas

Ies Dionisio Aguado

Estudiamos modelos de distribuciones probabilisticas por que su conocimiento como patrón de una serie de problemas similares tiene como resultado la simplificación del estudio de dichos problemas. El estudio de estos patrones simplificará el tratamiento estadístico de muchos fenómenos reales.

Por ejemplo, imaginemos que estudiamos un problema donde interviene la distribución binomial. Conociendo ésta distibución estariamos estudiando una abstracción matemática que permite identificar gran cantidad de problemas como el mismo problema.

1 Distribución uniforme discreta

La distribución uniforme discreta describe el comportamiento de una variable que puede tomar n valores distintos con la misma probabilidad cada uno de ellos. Por ejemplo, cuando se observa el número obtenido tras el lanzamiento de un dado perfecto, los valores posibles siguen una distribución uniforme discreta en $X = \{1, 2, 3, 4, 5, 6\}$, y la probabilidad de cada valor es $\frac{1}{6}$. Formalizando y generalizando el anterior ejemplo:

Sea X una variable aleatoria uniforme discreta, que toma los valores $x_1, \ldots, x_n$.

La **función de probabilidad** $f(x)$ viene dada por la siguiente expresión

$$f(x_i) = P(X = x_i) = \frac{1}{n} \quad i = 1, \ldots, n.$$

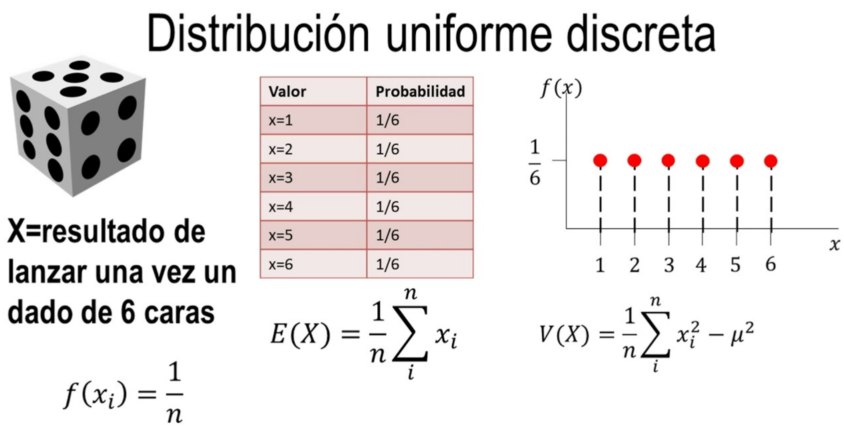
Suponiendo que $x_1 \leq x_2 \leq \ldots \leq x_n$, su **función de distribución** viene dada por

$$F(x) = \begin{cases} 0 & \text{si } x < x_1 \\ \frac{1}{n} & \text{si } x_1 \leq x < x_2 \\ \frac{2}{n} & \text{si } x_2 \leq x < x_3 \\ \vdots & \vdots \\ 1 & \text{si } x_n \leq x. \end{cases}$$

Por otro lado, se tiene que la esperanza o el valor esperado media poblacional o media de la distribución es $E[X] = \frac{1}{n}\sum_{i=1}^{n} x_i$ y que, $E[X^2] = \frac{1}{n}\sum_{i=1}^{n} x_i^2$.

Ejemplos

Para un dado perfecto, todos los resultados tienen la probabilidad de 1/6.



Para una moneda perfecta, todos los resultados tienen la probabilidad de 1/2.

2 Experimento de Bernouilli

Una ley de Bernoulli describe el comportamiento de un experimento aleatorio que tiene dos posibles resultados tradicionalmente llamados éxito y fracaso. Un experimento de este tipo se denomina ensayo de Bernoulli.

Consideremos la situación en la que un experimento aleatorio de este tipo (dos resultados posibles y una probabilidad fija) se repite un cierto número de veces de forma independiente; denotemos este número de veces. Esta repetición independiente de ensayos de Bernoulli se denomina Esquema de Bernoulli o simplemente Ensayos de Bernoulli.

Ejemplo

En un lanzamiento de cara o cruz, cara puede considerarse un éxito y cruz un fracaso. En este modelo, la probabilidad de éxito es un valor fijo, es decir, permanece constante en cada renovación del experimento aleatorio. Esta probabilidad de éxito se denomina "p".
Consideremos n lanzamientos sucesivos de una moneda. Estaremos ante un experimento de Bernuilli. Podríamos considerar «salir cara» como éxito y en este caso, la probabilidad de éxito es $p = \frac{1}{2}$y $q = \frac{1}{2}$

Obviamente se verifica que tanto p como q son estrictamente mayores que cero y además $p + q = 1$. Los dos posibles resultados suelen identificarse a menudo como *éxito* y *fracaso*, y la v.a. X que se asocia al experimento toma los valores 1 para el éxito y 0 para el fracaso.

2 Distribución binomial

Una **distribución binomial** es una distribución de probabilidad por la cual estudiamos las probabilidades de que haya k éxitos al realizar n pruebas Bernouilli independientes y con probabilidades de éxito iguales.

Ejemplo

El número de caras en 15 lanzamientos de una moneda sigue una distribución binomial.
Se consideran n variables aleatorias Bernouilli independientes, X_i con $i = 1, \ldots, n$, tal que, $X_i = 1$ si en el experimento Bernouilli se ha obtenido éxito y $X_i = 0$ en caso contrario. Obsérvese que si se define $X = \sum_{i=1}^{n} X_i$, se está contando el número de éxitos en n experimentos Bernouilli independientes. Por tanto, una v.a. con distribución binomial puede expresarse como suma de variables aleatorias independientes Bernouilli.
Así pues Si realizamos n veces un experimento en el que podemos obtener éxito, E, con probabilidad p y fracaso, F, con probabilidad $q(q = 1 - p)$, diremos que estamos ante una distribución binomial de parámetros n y p, y lo representaremos por $Bin(n; p)$.
En este caso la probabilidad de obtener k éxitos viene dada por:

$$P(X = k) = \binom{n}{k} p^k q^{n-k} \qquad k = 0, 1, 2, \ldots, n.$$

Para calcular la distribución de probabilidad de la variable binomial, se considera la ejecución n veces del experimento en el que han aparecido k éxitos y $n-k$ fracasos, –por ejemplo en la secuencia $EE \ldots EFF \ldots F$–, la probabilidad de este suceso viene dada por $p^k q^{n-k}$. Sin embargo, no es esa la única forma de obtener k éxitos, pues hay que considerar todas las ordenaciones de los elementos E y F, lo que da un total de:

$$P_n^{k,n-k} = \frac{n!}{k!(n-k)!} = \binom{n}{k}.$$

La probabilidad de obtener k éxitos en n pruebas es, por tanto:

$$P(X = k) = \binom{n}{k} p^k q^{n-k} \qquad k = 0, 1, 2, \ldots, n.$$

Será una verdadera distribución de probabilidad si la sma de probabilidades de los valores vale 1, y en efecto así es pues:

$$\sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} = (p + q)^n = 1^n = 1.$$

Por ejemplo $n = 3$, con probabilidad de éxito p: $B(3, p)$.

X	p	coeficientes
0	p^0q^3	$\binom{3}{0}$ =1
1	p^1q^2	$\binom{3}{1} = 3$
2	p^2q^1	$\binom{3}{2} = 3$
3	p^3q^0	$\binom{3}{3} = 1$

$$\sum_{k=0}^{3}\binom{3}{k}p^kq^{3-k} = \binom{3}{0}q^3 + \binom{3}{1}p^1q^2 + \binom{3}{2}p^2q^1 + \binom{3}{3}p^3$$
$$= 1q^3 + 3p^1q^2 + 3p^2q^1 + 1p^3 \qquad = (p+q)^3 = 1$$

De hecho, esta expresión justifica el nombre de la distribución, pues lo anterior no es sino el desarrollo del binomio $(p+q)^n$. A la distribución se le suele notar por $B(n,p)$, es decir por su letra o letras iniciales y los parámetros que la caracterizan. De igual forma se procede en el resto de los casos.

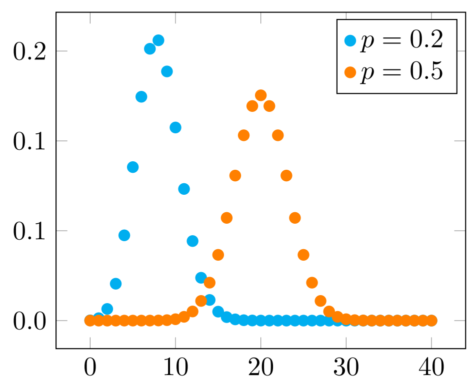
2.1.1 Parámetros de la distribución binomial

Los parámetros de una distribución binomial que necesitamos son su media, varianza y desviación típica.

$$media = E[X] = np,$$

$$varianza = \sigma^2 = V[X] = npq,$$

$$Desviación \quad típica = \sigma = \sqrt{npq}$$

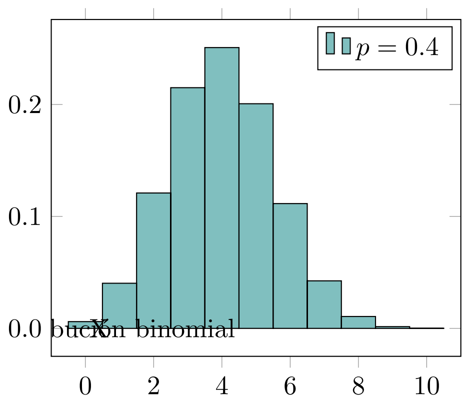


Ejemplo

De una población de animales, se sabe que el 60% son machos. Si se extrae un conjunto de 10 animales, ¿cuál es la probabilidad de que en ese conjunto haya 7 hembras?
Solución

Sea X = "Número de hembras en un conjunto de 10 animales",
Definamos las variables del experimento:
n = 10 (es el total de la muestra que tenemos)
x = número de éxitos, que en este caso es igual a 7, dado que buscamos la probabilidad de que 7 de los 10 sean hembras(éxito).
p = probabilidad de éxito (0.4)
q = probabilidad de fracaso (0.6)se ve claramente que $X \sim B(10, 0.4)$.
La probabilidad de que haya 7 hembras es:

$$P(X = 7) = \binom{10}{7} 0.4^7 0.6^3 = 0'04.$$



Ejemplo

Supongamos que se lanza 51 veces un dado de 6 caras y queremos calcular la probabilidad de que el número 3 salga 20 veces.
Solución
En este problema un ensayo consiste en lanzar el dado una vez. Consideramos un éxito si obtenemos un 3 pero si no sale 3 lo consideramos como un fracaso. Defínase X como el número de veces que se obtiene un «3» en 51 lanzamientos.
Definamos las variables del experimento:
n = 51 (es el total de la muestra que tenemos)
x = número de éxitos, que en este caso es igual a 20, dado que buscamos la probabilidad de que 20 de los 51 haya salido «3»(éxito).
p = probabilidad de éxito ($\frac{1}{6}$)
q = probabilidad de fracaso ($\frac{5}{6}$)
En este problema un ensayo consiste en lanzar el dado una vez. Consideramos un éxito si obtenemos un 3 pero si no sale 3 lo consideramos como un fracaso. Defínase X como el número de veces que se obtiene un «3» en 51 lanzamientos.

En este caso tenemos $X \sim Bin(51, 1/6)$ por lo que la probabilidad buscada es
$P[X = 20]$
$P[X = 20] = \binom{51}{20}(\frac{1}{6})^{20}(1 - \frac{1}{6})^{51-20} = 0.0000744$

Ejemplo

Imaginemos que un 80% de personas en el mundo ha visto el partido de la final del último mundial de fútbol. Tras el evento, 4 amigos se reúnen a conversar, ¿Cuál es la probabilidad de que 3 de ellos hayan visto el partido?
Solución
Definamos las variables del experimento:
n = 4 (es el total de la muestra que tenemos)
x = número de éxitos, que en este caso es igual a 3, dado que buscamos la probabilidad de que 3 de los 4 amigos lo hayan visto.
p = probabilidad de éxito (0,8)
q = probabilidad de fracaso (0,2). Este resultado se obtiene al restar 1-p.
Tras definir todas nuestras variables, simplemente sustituimos en la formula.
$P[X = 3] = \binom{4}{3}(0.8)^3(0.2)^{3-3} = 0.4096$
Si multiplicamos por 100 tenemos que hay una probabilidad del 40,96% de que 3 de los 4 amigos haya visto el partido de la final del mundial.

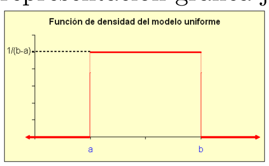
3 Distribución uniforme continua

A continuación se estudia la distribución uniforme continua como la extensión natural de la distribución uniforme discreta, es decir, aquella que toma con igual probabilidad valores dentro de dos conjuntos cualesquiera de igual amplitud e incluidos en el intervalo de valores posibles de la variable.

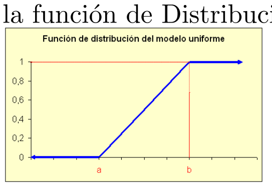
La variable X sigue una distribución uniforme o rectangular en el intervalo (a, b), $U(a, b)$, cuando su función de densidad viene dada de la siguiente forma:

$$f(x) = \begin{cases} \frac{x}{b-a} & \text{si} \quad a \leq x \leq b \\ 0 & \text{en el resto.} \end{cases}$$

Su representación gráfica justifica el nombre de rectangular



la función de Distribución será:



Sus características más importantes son:

$$E[X] = \frac{a+b}{2}, \quad V[X] = \frac{(b-a)^2}{12},$$

Ejemplo

Un autobús pasa por cierta parada cada 15 minutos. ¿Cuál es la probabilidad de que un señor que llega en un momento dado tenga que esperar el autobús más de cinco minutos?

Si se define X="Tiempo de espera", entonces $X \sim U(0, 15)$. Se calculará en primer lugar la función de distribución

$$F(x) = \begin{cases} 0 & \text{si} \quad x < 0 \\ \frac{x}{15} & \text{si} \quad 0 \leq x \leq 15 \\ 1 & \text{si} \quad x > 15. \end{cases}$$

La probabilidad pedida viene dada por:

$$P(X > 5) = 1 - P(X \leq 5) = 1 - \frac{5}{15} = 0'67.$$

4 Distribución normal

En este apartado se estudia la distribución más importante del cálculo de probabilidades y de la estadística, usualmente también denominada distribución de Gauss o de Laplace.

La importancia de esta distribución radica en varias razones:

La **distribución normal** es la distribución límite de una amplia gama de sucesiones de variables aleatorias independientes.

La gran mayoría de las variables aleatorias que se estudian en experimentos físicos son aproximadas por una distribución normal.

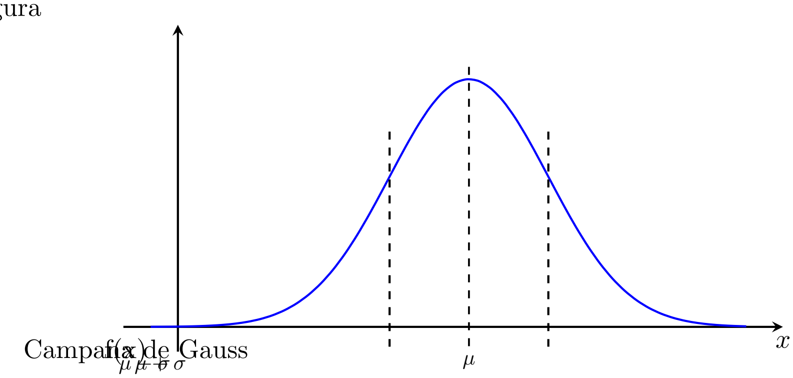
Los errores aleatorios en los resultados de medida se distribuyen, de forma general, según una distribución normal.

El carácter reproductivo de sus parámetros que facilita el manejo de este tipo de distribuciones.

Se dice que una variable X sigue una distribución normal si su función de densidad es:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad -\infty \leq x \leq +\infty.$$

La distribución está caracterizada por los parámetros μ y σ, cuyo significado se trata más adelante, siendo σ necesariamente positivo. Se hace referencia a esta distribución como $N(\mu, \sigma)$. Su representación gráfica viene dada por la figura



Algunos ejemplos de variables asociadas a fenómenos naturales que siguen el modelo de la normal son:

- caracteres morfológicos de individuos como la estatura;
- caracteres fisiológicos como el efecto de un fármaco;
- caracteres sociológicos como el consumo de cierto producto por un mismo grupo de individuos;
- caracteres psicológicos como el cociente intelectual;
- nivel de ruido en telecomunicaciones;
- errores cometidos al medir ciertas magnitudes; etc.

4 Distribución $N(0,1)$

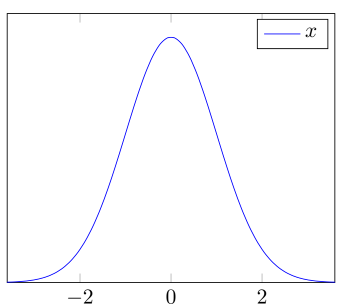
Para facilitar cálculos, se realiza el estudio de la distribución $N(0,1)$ y mediante un cambio de variable se generalizará para la $N(\mu,\sigma)$. La función de densidad de la $N(0,1)$ es:

$$f(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}} \qquad -\infty \leq x \leq +\infty.$$

A continuación se realiza el estudio de dicha función.

1. Campo de existencia. Toda la recta real para x, con $f(x)>0$.
2. Simetrías. Es simétrica respecto al eje OY, ya que $f(x)=f(-x)$.
3. Asíntotas. $y=0$ es una asíntota horizontal.
4. Cortes con los ejes. Sólo tiene un punto de corte en $(0,\frac{1}{\sqrt{2\pi}})$.
5. Máximos y mínimos. El punto de corte anterior es el único máximo de la distribución.
6. Puntos de inflexión. Tiene dos en $(-1,\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}})$ y $(1,\frac{1}{\sqrt{2\pi}}e^{\frac{-1}{2}})$.
7. El área de la región encerrada bajo la curva $f(x)$ es siempre 1.

Con todo lo anterior la función de densidad de una distribución normal tiene como gráfica la campana de Gauss



Las características de la distribución son:

$$\mu=E[X]=0, \quad \sigma^2=V[X]=1.$$

Los valores de la distribución $N(0,1)$ están tabulados y son fácilmente calculables.

A partir de ahora, si una variable sigue una distribución $N(0,1)$ se denotará por la letra Z. Así, si X sigue una $X=N(\mu,\sigma)$ se verifica que:

$$Z=\frac{X-\mu}{\sigma} \quad \text{y} \quad X=\sigma Z+\mu.$$

Con lo que se puede comprobar fácilmente que:

$$E[X] = \mu, \quad V[X] = \sigma^2.$$

Ejemplo

El contenido de un tipo de bombonas de gas se distribuye normalmente con media 23 kg y desviación típica 0.25 kg. ¿Cuál es la probabilidad de que una bombona tenga más de 23.5 Kg?
La probabilidad pedida es:

$$P(X > 23'5) = 1 - P(X \leq 23'5).$$

Para calcular esta probabilidad hay que tipificar la variable X y hacer uso de la tabla $N(0,1)$

$$\begin{aligned} P(X \leq 23'5) &= P\left(\frac{X-\mu}{\sigma} \leq \frac{23.5-\mu}{\sigma}\right) \\ &= P\left(Z \leq \frac{23.5-23}{0.25}\right) \\ &= P(Z \leq 2) = 0 - 0.977. \end{aligned}$$

Por lo tanto:

$$P(X > 23.5) = 1 - 0.977 = 0.023.$$

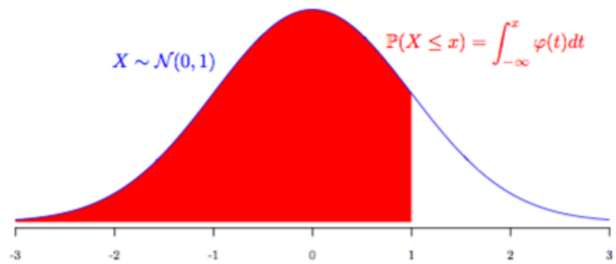
5 Uso de la tabla de la Normal $Z = N(0,1)$

La tabla con la que vamos a trabajar nos da las probabilidades $P[Z \leq x]$ para valores de x de 0 a 4. A estas probabilidades se les llama

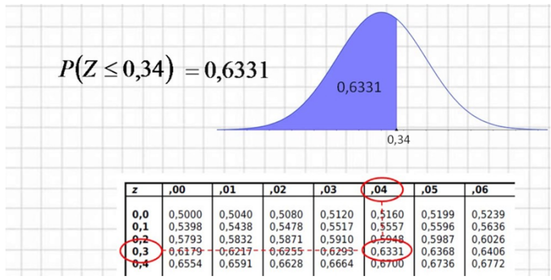
$$F(x) : F(x) = P[Z \leq x]$$

Z se distribuye según una $N(0,1)$.$F(x)$ es, pues, la función de distribución de esta variable aleatoria.

Esta tabla de doble entrada, presenta la probabilidad para $Z < x$, de la distribución normal acumulada, para valores de x iguales o mayores que cero, en la columna izquierda esta la parte entera y el primer decimal de x, y en la fila superior el segundo decimal del número, en la casilla donde se cruzan la fila y la columna correspondientes, figura el valor de la probabilidad de que Z < x, con varias cifras decimales(es decir el área acumulada desde $-\infty$ hasta el valor x.

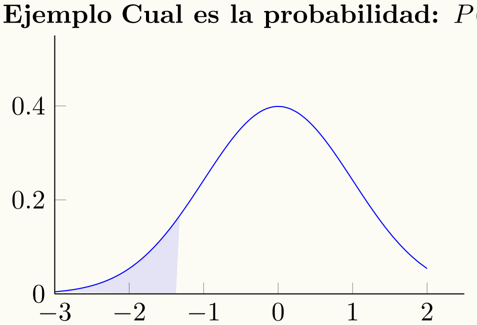


	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

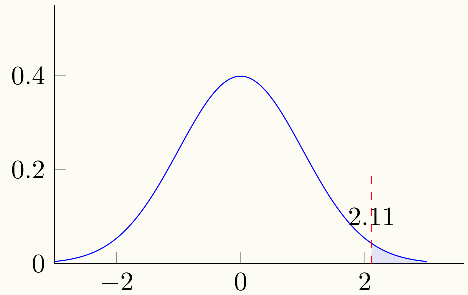


z	,00	,01	,02	,03	,04	,05	,06
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772

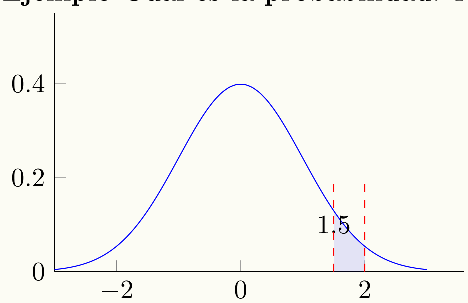
Ejemplos
Ejemplo Buscar la probabilidad normal acumulada de que $Z < 2,04$.
En la columna del 2 y la fila del 0,04, esta el valor 0,979324, esto es: $P(Z_{(0,1)} < 2,04) = 0,979324$ ó 97.93% de probabilidad.
En la tabla anterior se pueden buscar los valores de la probabilidad normal tipificada:
$P(Z_{(0,1)} < x)$ para valores de x mayores o iguales a cero, como el ejemplo anterior, hay más casos, que con los datos de la tabla se pueden resolver.
Para x < 0
Ejemplo Cual es la probabilidad: $P(Z_{(0,1)} < -1,32)$



$P(Z_{(0,1)} < -1,32)$ los valores negativos no vienen en la tabla, entonces:
$P(Z_{(0,1)} < -1,32) = 1 - P(Z_{(0,1)} < 1,32)$
$P(Z_{(0,1)} < -1,32) = 1 - P(Z_{(0,1)} < 1,32)$ según la tabla:
$P(Z_{(0,1)} < 1,32) = 0,906582$ por tanto:
$P(Z_{(0,1)} < -1,32) = 1 - 0,906582$ que resulta:
$P(Z_{(0,1)} < -1,32) = 0,093417$
Ejemplo: Cual es la probabilidad: $P(Z_{(0,1)} > 2,11)$



$P(Z_{(0,1)} > 2,11) = 1 - P(Z_{(0,1)} < 2,11)$ de la tabla tenemos:
$P(Z_{(0,1)} < 2,11) = 0,982570$ lo que resulta:
$P(Z_{(0,1)} > 2,11) = 1 - 0,982570$ que resulta:
$P(Z_{(0,1)} > 2,11) = 0,017430$
Ejemplo Cual es la probabilidad: $P(Z_{(0,1)} > -2,02)$
$P(Z_{(0,1)} > -2,02) = P(Z_{(0,1)} < 2,02)$ buscando el valor en la tabla, tenemos que:
$P(Z_{(0,1)} > -2,02) = P(Z_{(0,1)} < 2,02) = 0,978308$
Ejemplo Cual es la probabilidad: $P(1,50 < Z_{(0,1)} < 2,00)$



se buscan en la tabla las probabilidades:
$P(Z_{(0,1)} < 1,50) = 0,933192$ $P(Z_{(0,1)} < 2,00) = 0,977249$
Luego:
$P(1,50 < Z_{(0,1)} < 2,00) = P(Z_{(0,1)} < 2.00) - P(Z_{(0,1)} < 1,50)$ esto es:
$P(1,50 < Z_{(0,1)} < 2,00) = 0,977249 - 0,933192$
Realizando la operación:
$P(1,50 < Z_{(0,1)} < 2,00) = 0,044057$

5 Resumen uso solución ejercicios de la Normal

Sea Z=N(0,1) la variable aleatoria Normal de parámetros : media $\mu = 0$ y desviación típica $\sigma = 1$

Entonces si se usa una tabla del area acumulada a la izquierda de los valores a>0

- P(Z=a)=0

- $P(Z \leq a) = P(Z < a)$= Mirar directamente en la tabla con el valor de a
- P(Z≤a)=P(Z>a)= 1-valor en Tabla(a)
- P(Z≤ $-a$)=P(Z<-a)= 1-valor en Tabla(a)
- P(Z≥-a)=P(Z>-a)= valor en Tabla(a)
- P(a≤Z≤b)=P(a<Z<b)=P(a<Z≤b)=P(a≤Z<b)=valor en Tabla(b) - valor enTabla (a)
- P(-a≤Z≤b)=P(-a<Z<b)=P(-a<Z≤b)=P(-a≤Z<b)=valor en Tabla(b)+valor en Tabla (a)-1
- P(-a≤Z≤ $-$b)=P(-a<Z< $-$b)=P(-a<Z≤-b)=P(-a≤Z< $-$b)=valor en Tabla(a)+valor en Tabla (b)

Si la variable aleatoria X es una Normal de parámetros μ y σ,es decir ,$X \sim N(\mu,\sigma)$, para calcular probabilidades se hará un cambio de variable$Z \sim \frac{X-\mu}{\sigma}$ que convierte la variable X en una variable $Z \sim N(0,1)$.

El cambio de variable se denomina Tipificar la variable

Para ilustrar el tema supongamos que $X \sim N(\mu,\sigma)$ y queremos calcular $P(X \geq a) = P(Z \geq \frac{a-\mu}{\sigma})$, esta última probabilidad se calculará con ayuda de la tabla de la Normal $Z \sim N(0,1)$ según el valor de $\frac{a-\mu}{\sigma}$.

Ejemplo

Si X es una variable aleatoria de una distribución $N(\mu,\sigma)$, hallar:
$P(\mu - 3\sigma \leq X \leq \mu + 3\sigma)$.

Solución:

sabemos que = $Z = \frac{X-\mu}{\sigma}$
; como $x_1 = \mu - 3\sigma$ y $x_2 = \mu + 3\sigma$ entonces la $P(x_1 \leq X \leq x_2)$ se define como:

$$P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) = P(\frac{(\mu-3\sigma)-\mu}{\sigma} \leq X \leq \frac{\mu + 3\sigma - \mu}{\sigma})$$

$$= P(-3 \leq Z \leq 3) = P(Z \leq 3) - P(Z \leq -3)$$

$$= P(-3 \leq Z \leq 3) = P(Z \leq 3) - (1 - P(Z \leq 3)) = 0.9986 - 1 + 0.9986 = 0.9972$$

Es decir, que aproximadamente el 99.72% de los valores de X están a más/menos de tres desviaciones típicas de la media.

En una ciudad se estima que la temperatura máxima en el mes de junio sigue una distribución normal, con media 23° y desviación típica 5°. Calcular el número de días del mes en los que se espera alcanzar máximas entre 21° y 27°.
Solución: Sabemos que $Z = \frac{X-\mu}{\sigma}$, como $x_1 = 21$ y $x_2 = 27$ entonces la $P(x_1 \leq X \leq x_2)$ se define como:

$$P(21 \leq X \leq 27) = P(\frac{(21)\text{-}\mu}{\sigma} \leq X \leq \frac{27-\mu}{\sigma})$$

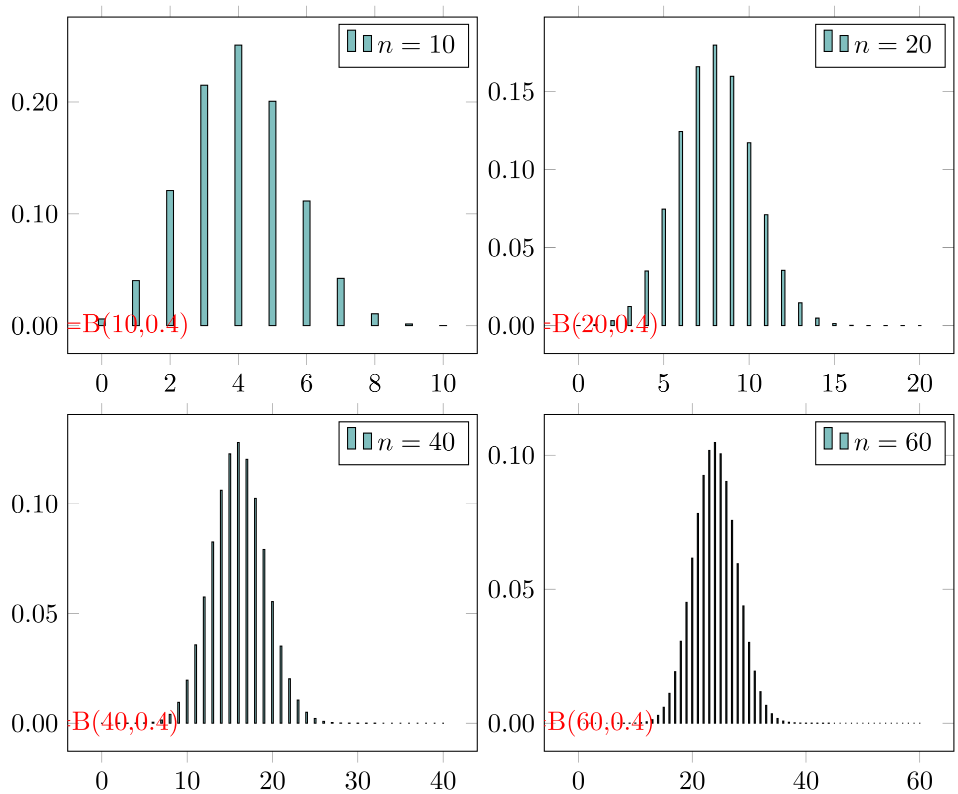
$$= P(\frac{21-23}{5} \leq Z \leq \frac{27-23}{5}) = P(\frac{-2}{5} \leq Z \leq \frac{4}{5})$$

$$= P(-0.4 \leq Z \leq 0.8) = P(Z \leq 0.8) - (1 - P(Z \leq 0.4)) =$$

$$0.7881 + 0.6554 - 1 = 0.4435; \quad 0.4435 * 30 = 13.3 \quad días$$

Binomial aproximada a la normal

Vamos a representar en un sistema de referencia distribuciones binomiales para distintos valores de n y $p = 0,4$.



Se puede apreciar en los gráficos anteriores como a medida que aumenta n mejora el parecido de las gráficas de barras de las distribuciones binomiales (discretas) a la gráfica de la distribución normal estándar (continua),.

Sea X una variable aleatoria discreta tal que $X \sim B(n,p)$. Cuando n es suficientemente grande, la distribución binomial se puede aproximar a una distribución normal de media $\mu = np$ y desviación estándar $\sigma = \sqrt{npq}$

Para que $X \sim N(np, \sqrt{npq})$ sea una buena aproximación de $X \sim B(n,p)$ debe cumplirse que:

$$n \geq 30; \;\; np > 5 \;\; n(1-p) > 5$$

Ejemplo

La probabilidad de dar en la diana al lanzar una flecha con arco es 0,75. Si lanzamos 100 veces una flecha, ¿Cuál es la probabilidad de hacer 77 dianas o más?

Solución:

Se trata de un distribución binomial B(100; 0,75) y debemos calcular $P(X \geq 77)$. Los cálculos son excesivos ya que la solución es
$P(X = 77) + P(X = 78) + ... + P(X = 100)$
Intentamos ver si es posible resolverlo por la distribución Normal
$n = 100, p = 0.75,\ q = 1 - 0.75 = 0.25$
$np = 75$ y $nq = 25$ así que se cumplen las condiciones. Resolvemos usando la distribución Normal.
B(100.0.75) pasa a ser aproximada por una Normal donde la media es $\mu = np = 75$ y la desviación es $\sigma = \sqrt{npq} = \sqrt{100 \cdot 0.75 \cdot 0.25} = 4.3$ $\sim N(75, 4.3)$
$P(X \geq 77) = P(X' > 76,5)$ =$P(Z > \frac{76.5-75}{4.3}$)= $P(Z > 0.35)$=1$-P(Z <$ 0.35)=0.6368
Usamos 76.5 ya que en el paso de una distribución discreta(solo naturales) a una continua (todos los números) el intervalo [76,77] debe ser repartido. Así si nos referimos al valor 77 en la discreta, nos referimos al intervalo [76.5,77.5] en la continua.

6 Ejercicios propuestos

Ejercicios de distribución binomial

1. Se lanza una moneda cuatro veces. Calcular la probabilidad de que salgan más caras que cruces.

2. Un agente de seguros vende pólizas a cinco personas de la misma edad y que disfrutan de buena salud. Según las tablas actuales, la probabilidad de que una persona en estas condiciones viva 30 años o más es 2/3. Hállese la probabilidad de que, transcurridos 30 años, vivan: 1. Las cinco personas. 2. Al menos tres personas. 3. Exactamente dos personas.

3. Si de seis a siete de la tarde se admite que un número de teléfono de cada cinco está comunicando, ¿cuál es la probabilidad de que, cuando se marquen 10 números de teléfono elegidos al azar, sólo comuniquen dos?

4. La probabilidad de que un hombre acierte en el blanco es 1/4. Si dispara 10 veces ¿cuál es la probabilidad de que acierte exactamente en tres ocasiones? ¿Cuál es la probabilidad de que acierte por lo menos en una ocasión?

5. En una urna hay 30 bolas, 10 rojas y el resto blancas. Se elige una bola al azar y se anota si es roja; el proceso se repite, devolviendo la bola, 10 veces. Calcular la media y la desviación típica.

6. La última novela de un autor ha tenido un gran éxito, hasta el punto de que el 80% de los lectores ya la han leido. Un grupo de 4 amigos son aficionados a la lectura: 1. ¿Cuál es la probabilidad de que en el grupo hayan leido la novela 2 personas? 2. ¿Y cómo máximo 2?

7. Un agente de seguros vende pólizas a cinco personas de la misma edad y que disfrutan de buena salud. Según las tablas actuales, la probabilidad de que una persona en estas condiciones viva 30 años o más es 2/3. Hállese la probabilidad de que, transcurridos 30 años, vivan: 1. Las cinco personas. 2. Al menos tres personas. 3. Exactamente dos personas.

8. Se lanza una moneda cuatro veces. Calcular la probabilidad de que salgan más caras que cruces.

9. Si de seis a siete de la tarde se admite que un número de teléfono de cada cinco está comunicando, ¿cuál es la probabilidad de que, cuando se marquen 10 números de teléfono elegidos al azar, sólo comuniquen dos? 5 La probabilidad de que un hombre acierte en el blanco es 1/4. Si dispara 10 veces ¿cuál es la probabilidad de que acierte exactamente en tres ocasiones? ¿Cuál es la probabilidad de que acierte por lo menos en una ocasión?

10. En unas pruebas de alcoholemia se ha observado que el 5% de los conductores controlados dan positivo en la prueba y que el 10% de los conductores controlados no llevan aprovechado el cinturón de seguridad. También se ha observado que las dos infracciones son independientes. Un guardia de tráfico para cinco conductores al azar. Si tenemos en cuenta que el número de conductores es suficientemente importante como para estimar que la proporción de infractores no varía al hacer la selección. 1. Determinar la probabilidad a de que exactamente tres conductores hayan cometido alguna de las dos infracciones. 2. Determine la probabilidad de que al menos uno de los conductores controlados haya cometido alguna de las dos infracciones.

11. La probabilidad de que un artículo producido por una fabrica sea defectuoso es p 0.002. Se envió un cargamento de 10.000 artículos a unos almacenes. Hallar el número esperado de artículos defectuosos, la varianza y la desviación típica.

12. En una urna hay 30 bolas, 10 rojas y el resto blancas. Se elige una bola al azar y se anota si es roja; el proceso se repite, devolviendo la bola, 10 veces. Calcular la media y la desviación típica.

13. Un laboratorio afirma que una droga causa de efectos secundarios en una proporción de 3 de cada 100 pacientes. Para contrastar esta afirmación, otro laboratorio elige al azar a 5 pacientes a los que aplica la droga. ¿Cuál es la probabilidad de los siguientes sucesos?

    (a) Ningún paciente tenga efectos secundarios.

    (b) Al menos dos tengan efectos secundarios.

    (c) ¿Cuál es el número medio de pacientes que espera laboratorio que sufran efectos secundarios si elige 100 pacientes al azar?

EJERCICIOS SOBRE UNIFORME Y NORMAL

1. . El salario mensual de un vendedor se distribuye de manera uniforme entre $120.00, que es su sueldo base, y $780.00.

    (a) ¿Cuál es la probabilidad de que en un mes cualquiera gane más de $550.00?

    (b) ¿ Y de que gane menos de 700? **sol:(0.348, 0.879)**

2. El precio medio del galón de gasolina regular durante el próximo año se estima que puede oscilar entre 4.10 y 4.50 de dólar. Si para tramos iguales de precios dentro del intervalo se tienen iguales probabilidades, determinar:

    (a) probabilidad de que el precio sea mayor que 4.20

    (b) se sitúe entre 4.22 y 4.36. **sol:(0.75; 0.35)**

3. El tiempo de vuelo en minutos de un avión entre dos ciudades A y B, puede ser cualquier valor entre 125 y 145 minutos. Supóngase que la probabilidad de un tiempo de vuelo dentro de un intervalo específico, es la misma para cualquier otro intervalo de la misma longitud; dentro del intervalo de tiempo de vuelo. Determinar las probabilidades de que, el tiempo de vuelo:
   (a) sea mayor de 132 minutos
   (b) dure entre 130 y 138 minutos. **sol(0.65; 0.40)**
4. Dada N (74, 144), hallar los puntajes estándar que corresponden a:
   (a) 65 ,
   (b) 74 ,
   (c) 86. **sOL(0.75; 0; 1)**
5. Dada N (74, 144), hallar el valor de la variable que corresponde a los puntajes:
   (a) z = -1,
   (b) z = 1.25,
   (c) z = 1.82 **Sol(62; 89; 95.84)**
6. Hallar el área correspondiente a los sucesos en la curva normal estándar y representarla gráficamente:
   (a) z > 1.22
   (b) -2 < z <2
   (c) z <-0.5 **Sol(0.1112; 0.9544; 0.3085)**
7. Una máquina debe llenar botes de jalea con gramos de producto, pero no lo hace exactamente. Los pesos reales de contenido siguen una ley normal.
   (a) a) Si μ=300 y σ=4, ¿qué proporción de botes tendrán menos de 308 gramos?.
   (b) Si la desviación se mantiene en 4 gramos, ¿a qué valor debe ajustarse la media para que solamente 2% de los botes contengan menos de 300 gramos de producto?. **Sol(0.0228, 308.2)**
8. Las calificaciones finales de un curso con 120 estudiantes siguen una distribución normal, con media de 6.5 y una varianza de 1.21.
   (a) Si la menor nota para aprobar el curso es de 6, ¿cuántos alumnos reprobaron la materia?
   (b) ¿Qué calificación era necesaria sobrepasar para formar parte del 10% con mejor rendimiento? **sol(32, 7.9)**

9. Los resultados de un curso de adiestramiento para los obreros de una fábrica se aproximan a una ley normal con μ=76.2 y σ=3. Se ha decidido que el 22% de aquellos que obtuvieron las menores notas continúen un adiestramiento complementario. ¿ Cuál es la nota mínima para no hacer el curso complementario?. sol(70)

10. El tiempo en minutos que los aspirantes a un puesto de trabajo tardan en realizar una prueba de conocimientos generales, sigue una ley normal con media 68 y varianza 64.

    (a) Qué proporción de los aspirantes tarda entre 60 y 80 minutos? **sol(77.45%)**

    (b) El 15% de los que tardaron menos tiempo en realizar el examen, harán una prueba psicológica. ¿Cuál es el tiempo mínimo de entrega del examen de conocimientos, que da la oportunidad de realizar la prueba psicológica? **sol(59.68 minutos)**

11. Se ha estimado que el tiempo de espera en una sucursal bancaria, sigue aproximadamente una ley normal de probabilidad con media de 7.8 minutos y desviación estándar de 1.6 minutos. Si se elige un cliente de manera aleatoria, determínese la probabilidad de que:

    (a) Haya esperado más de 10 minutos **sol(0.0838)**

    (b) Haya esperado entre 6 y 9 minutos **sol(0.6442)**

12. El régimen de lluvia anual en pulgadas, en cierta región del país sigue una ley normal μ=30 y σ=8.1 con y , hallar las probabilidades de lluvia anual:

    (a) mayor que 33 pulgadas (0.1587)

    (b) entre 27 y 32 (0.5899)

13. En una droguería se deben de llenar botellas de alcohol con 250 mililitros. Supóngase que los volúmenes de llenado se distribuyen aproximadamente normal y que la desviación estándar se fija en 4 ml.

    (a) Determinar la probabilidad de que una botella seleccionada al azar, tenga menos de 245 ml. **sol(0.1056)**

    (b) l supervisor de control de calidad rechaza aquellas botellas que tiene más de 5 ml de diferencia respecto de la media, ¿ cuál es la probabilidad de rechazo? **sol(0.2112)**

14. Se ha comprobado que el tiempo que tardan las personas en llenar un formulario de el Ministerio de Hacienda sigue una distribución normal con media de 42 minutos y desviación estándar de 8 minutos.

    (a) Si se elige una persona al azar, ¿cuál es la probabilidad que tarde entre 40 y 50 minutos en llenar el formulario? (0.44)

    (b) ¿Cuál es la probabilidad que tarde menos de 30 minutos? (0.0668)

15. (Aproximación de binomial por normal). Si en una distribución binomial la probabilidad p de éxito es próxima a 0.5 y se tiene np $>$ 0.5 y nq $>$ 0.5, entonces la variable: se comporta aproximadamente como una distribución normal $N(np, \sqrt{npq})$. Y el problema binomial se puede resolver empleando la distribución normal. Si X es binomial $B(100, 0.45)$, calcular P(20$<$ x $<$ 70)

    (a) usando binomial ,

    (b) usando la aproximación por normal.