Distribución de Probabilidad

Departamento de Matemáticas
IES Pedro Cerrada

1 Variables aleatorias y distribuciones de probabilidad

Finalidad : Abstraer matemáticamente un tipo de experimento aleatorio. Y, con ello, poder estimar de manera teórica lo que sucedería de manera experimental mediante una estadística.

¿Cómo?: Mediante variables aleatorias y distribuciones de probabilidad asociadas a esas variables:

2 Variables aleatorias

Una variable aleatoria es una función que a cada suceso elemental de un espacio muestral le asigna un número. Para hacer referencia a las variables se usan las letras: X, Y, ...

Ejemplo: Sea el experimento aleatorio "lanzar un dado": El espacio muestral lo componen las 6 caras del dado. Podemos asignar la variable X que a cada cara le asocia el número que represente su cara.

Suceso	X	x_i
Cara 1	$\rightarrow$	1
Cara 2	$\rightarrow$	2
Cara 3	$\rightarrow$	3
Cara 4	$\rightarrow$	4
Cara 5	$\rightarrow$	5
Cara 6	$\rightarrow$	6

Ejemplo: Sea el experimento compuesto lanzar dos monedas. Podemos asignar la variable aleatoria:

Y = *{Número de caras}*

Suceso	Y	y_i
C,C	$\rightarrow$	2
C,X	$\rightarrow$	1
X,C	$\rightarrow$	1
X,X	$\rightarrow$	0

2.1. Tipos de variables aleatorias

Discretas: Toman un número finito o numerable de valores

Continuas: Toman valores en un rango continuo

Ejemplo de variables discreta: Sea la X =*"El número de caras al lanzar dos dados"*. Los valores posibles son 0, 1 o 2 (que es un conjunto finito de datos, en concreto 3 datos)

Ejemplo de variables continua: X = *"Distancia al centro de la diana medida desde la posición en que cae un dardo lanzado por un tirador experto"* . En este caso la variable puede tomar cualquier valor en el rango entre 0 y el radio de la diana

3 Distribuciones de probabilidad

Llamaremos Distribución de probabilidad a la relación entre los valores de la variable y sus probabilidades.

Estas relaciones se pueden indicar mediante el uso de funciones. El tipo de función y su tratamiento es diferente según las variables sean discretas o continuas.

Veamos algunas distribuciones:

4 Distribución uniforme discreta

Ejemplo: Sea la variable X = *"Número obtenido al lanzar una dado"*
A cada valor de la variable podemos asignarle su probabilidad:

x_i	$P(X)$	$P(x_i)$
1	$\rightarrow$	$\frac{1}{6}$
2	$\rightarrow$	$\frac{1}{6}$
3	$\rightarrow$	$\frac{1}{6}$
4	$\rightarrow$	$\frac{1}{6}$
5	$\rightarrow$	$\frac{1}{6}$
6	$\rightarrow$	$\frac{1}{6}$

Podemos representar la relación anterior mediante una función:

$$P: \begin{array}{l} X \rightarrow \mathbb{R} \\ x_i \mapsto P(X = x_i) = \dfrac{1}{n} \end{array}$$

Todas las caras tienen la misma probabilidad: $\frac{1}{n}$, siendo n el número de caras, o tamaño del espacio muestral.

A este tipo de distribución se le llama **uniforme discreta**.

5 Distribución Binomial

Ejemplo: Queremos calcular las probabilidades de que al lanzar 5 monedas, obtengamos tres caras. Un suceso que cumple el enunciado es:

$$S_1 = \{C, C, C, X, X\}$$

Teniendo en cuenta que lanzar cada moneda son experimentos independientes, la probabilidad de ese suceso será:

$$\begin{aligned} P(S_1) &= P(C_1)\cdot P(C_2|C_1)\cdot ...\cdot P(X_5|C_1\cap C_2\cap C_3\cap X_4) = \\ &= P(C_1)\cdot P(C_2)\cdot P(C_3)\cdot P(X_4)\cdot P(X_5) = \\ &= P(C)^3\cdot P(X)^2 \end{aligned}$$

Como la probabilidad de que una moneda sea cara es $\frac{1}{2}$ y la de que sea cruz también:

$$P(S_1) = \left(\frac{1}{2}\right)^3\cdot\left(\frac{1}{2}\right)^2$$

Ahora bien, habrá tantos sucesos que cumplan el enunciado como combinaciones de 5 elementos tomados de 3 en 3. Por tanto la probabilidad de que salgan 3 caras será:

$$P(Salgan\ 3\ caras) = \binom{5}{3}\left(\frac{1}{2}\right)^3\cdot\left(\frac{1}{2}\right)^2$$

Si asociamos al experimento "lanzar 5 monedas"le asignamos la variable "número de caras obtenidas", podemos determinar la probabilidad mediante la siguiente función.

$$P:\begin{array}{l} X\to\mathbb{R} \\ k\mapsto P(X=k) = \binom{5}{k}\left(\frac{1}{2}\right)^k\cdot\left(\frac{1}{2}\right)^{5-k}\end{array}$$

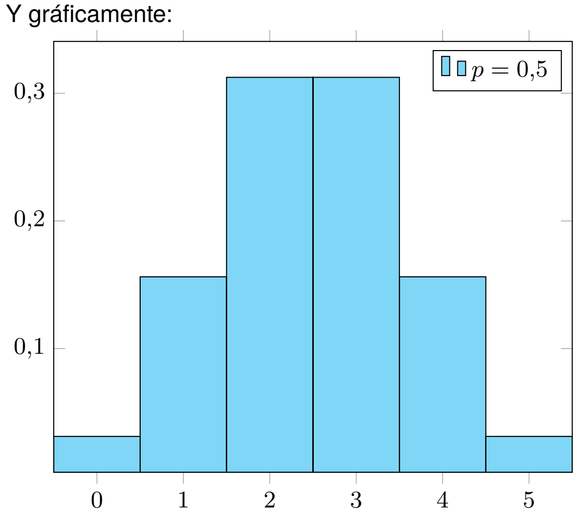
Esto es un ejemplo de distribución binomial de tamaño 5 y probabilidad $0{,}5$.

Realizando los cálculos para $k = 1,...,5$, la distribución de probabilidad de X que resulta es:

$P(X)$

x_i		$P(x_i)$
0	$\to$	0,03125
1	$\to$	0,15625
2	$\to$	0,3125
3	$\to$	0,3125
4	$\to$	0,15625
5	$\to$	0,03125

Y gráficamente:



Generalización de distribución Binomial: Hablaremos de una distribución binomial cuando:

- Se parte de un experimento compuesto de varios simples independientes
- Los experimentos simples son dicotómicos. Es decir, solo puede haber dos sucesos elementales: uno al que llamaremos acierto y otro al que llamaremos fracaso
- Asociado al experimento compuesto tenemos la variable número de aciertos cuando realizamos el experimento simple un número determinado de veces

En la situación anterior, la distribución binomial vendrá determinada por dos parámetros:

- Parámetro n: Número de veces que se realiza el experimento simple
- Parámetro p: La probabilidad de que ocurra el suceso acierto

A este tipo de variable y su distribución de probabilidades se le llama binomial.

Ejemplo: Así en el ejemplo de las monedas, el experimento se compone de **5 lanzamientos de moneda**. Si sale cara es acierto y si no fracaso. La variable aleatoria asociada al experimento será el número de caras que salen al lanzar 5 monedas. Esta variable sigue una distribución binomial.

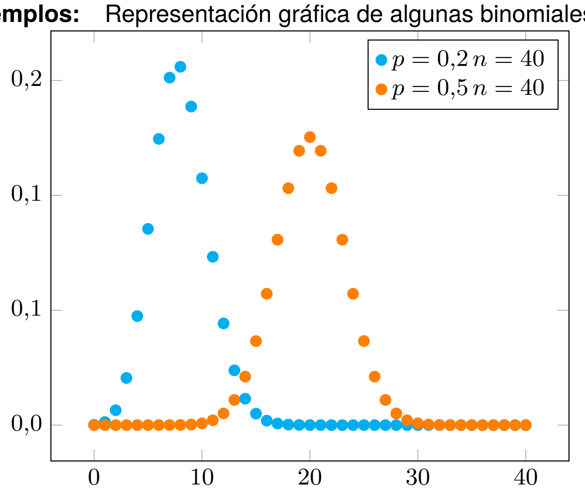
5.1. Probabilidad de la binomial

En general tendremos una binomial de tamaño n y probabilidad p, cuando el experimento simple se haga n veces y la probabilidad de acierto sea p.

La función de probabilidad en este caso nos queda:

$$P:\begin{array}{l} X\to\mathbb{R} \\ k\mapsto P(X=k) = \binom{n}{k}(p)^k\cdot(1-p)^{n-k}\end{array}$$

Ejemplos: Representación gráfica de algunas binomiales:



6 Distribución Normal

Se trata de una distribución asociada a a un variable continua. Tiene dos parámetros:

- μ : Media de la distribución
- σ : Desviación típica

Y se denota así: $X \sim \mathcal{N}(\mu, \sigma)$.

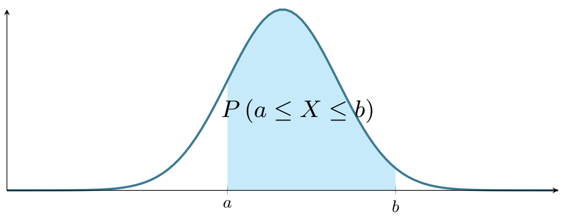
La probabilidad total es 1 y la variable X puede tomar ∞ valores ($x \in (-\infty, +\infty)$). Por tanto, en variables continuas la probabilidad de que tome un valor concreto será 0 ($\lim\limits_{x\to\infty} \frac{1}{x} = 0$).

6.1. función de densidad de una variable continua:

En variables continuas solo tiene sentido calcular la probabilidad en intervalos. Se llama **función de densidad** ($f(x)$) a aquella que:

$$P(a \leq X \leq b) = \int_a^b f(x)dx$$

La interpretación gráfica de lo anterior nos dice que la probabilidad de un intervalo corresponde con el área de la función de densidad en ese intervalo.

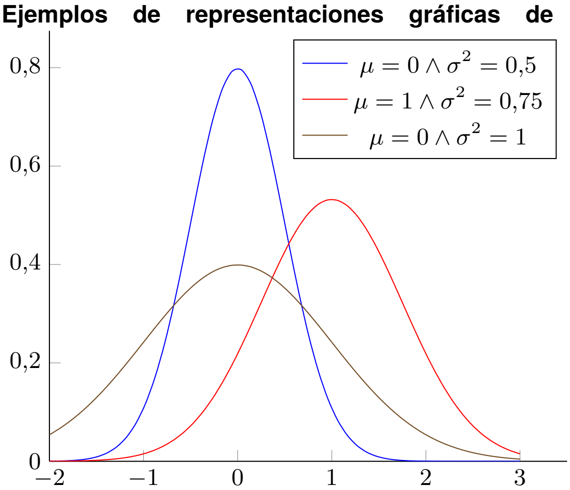


6.2. función de densidad de una distribución normal:

En el caso de una $X \sim \mathcal{N}(\mu, \sigma)$, la función de densidad es:

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

Ejemplos de representaciones gráficas de normales:



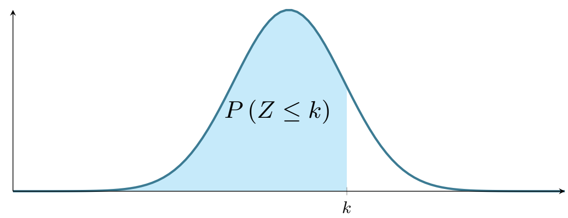
6.3. Cálculo práctico de la probabilidad de la Normal:

En realidad, para calcular la probabilidad no se hace la integral, sino que se utiliza una tabla que ya tiene calculadas probabilidades de la $Z \sim \mathcal{N}(0, 1)$:

z	0	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0	0,5	0,50399	0,50798	0,51197	0,51595	0,51994	0,52392	0,5279	0,53188	0,53586
0,1	0,53983	0,5438	0,54776	0,55172	0,55567	0,55962	0,56356	0,56749	0,57142	0,57535
0,2	0,57926	0,58317	0,58706	0,59095	0,59483	0,59871	0,60257	0,60642	0,61026	0,61409
0,3	0,61791	0,62172	0,62552	0,6293	0,63307	0,63683	0,64058	0,64431	0,64803	0,65173
0,4	0,65542	0,6591	0,66276	0,6664	0,67003	0,67364	0,67724	0,68082	0,68439	0,68793
0,5	0,69146	0,69497	0,69847	0,70194	0,7054	0,70884	0,71226	0,71566	0,71904	0,7224
0,6	0,72575	0,72907	0,73237	0,73565	0,73891	0,74215	0,74537	0,74857	0,75175	0,7549
0,7	0,75804	0,76115	0,76424	0,7673	0,77035	0,77337	0,77637	0,77935	0,7823	0,78524
0,8	0,78814	0,79103	0,79389	0,79673	0,79955	0,80234	0,80511	0,80785	0,81057	0,81327
0,9	0,81594	0,81859	0,82121	0,82381	0,82639	0,82894	0,83147	0,83398	0,83646	0,83891
1	0,84134	0,84375	0,84614	0,84849	0,85083	0,85314	0,85543	0,85769	0,85993	0,86214
1,1	0,86433	0,8665	0,86864	0,87076	0,87286	0,87493	0,87698	0,879	0,881	0,88298
1,2	0,88493	0,88686	0,88877	0,89065	0,89251	0,89435	0,89617	0,89796	0,89973	0,90147
1,3	0,9032	0,9049	0,90658	0,90824	0,90988	0,91149	0,91309	0,91466	0,91621	0,91774
1,4	0,91924	0,92073	0,9222	0,92364	0,92507	0,92647	0,92785	0,92922	0,93056	0,93189
1,5	0,93319	0,93448	0,93574	0,93699	0,93822	0,93943	0,94062	0,94179	0,94295	0,94408
1,6	0,9452	0,9463	0,94738	0,94845	0,9495	0,95053	0,95154	0,95254	0,95352	0,95449
1,7	0,95543	0,95637	0,95728	0,95818	0,95907	0,95994	0,9608	0,96164	0,96246	0,96327
1,8	0,96407	0,96485	0,96562	0,96638	0,96712	0,96784	0,96856	0,96926	0,96995	0,97062
1,9	0,97128	0,97193	0,97257	0,9732	0,97381	0,97441	0,975	0,97558	0,97615	0,9767
2	0,97725	0,97778	0,97831	0,97882	0,97932	0,97982	0,9803	0,98077	0,98124	0,98169
2,1	0,98214	0,98257	0,983	0,98341	0,98382	0,98422	0,98461	0,985	0,98537	0,98574
2,2	0,9861	0,98645	0,98679	0,98713	0,98745	0,98778	0,98809	0,9884	0,9887	0,98899
2,3	0,98928	0,98956	0,98983	0,9901	0,99036	0,99061	0,99086	0,99111	0,99134	0,99158
2,4	0,9918	0,99202	0,99224	0,99245	0,99266	0,99286	0,99305	0,99324	0,99343	0,99361
2,5	0,99379	0,99396	0,99413	0,9943	0,99446	0,99461	0,99477	0,99492	0,99506	0,9952
2,6	0,99534	0,99547	0,9956	0,99573	0,99585	0,99598	0,99609	0,99621	0,99632	0,99643
2,7	0,99653	0,99664	0,99674	0,99683	0,99693	0,99702	0,99711	0,9972	0,99728	0,99736
2,8	0,99744	0,99752	0,9976	0,99767	0,99774	0,99781	0,99788	0,99795	0,99801	0,99807
2,9	0,99813	0,99819	0,99825	0,99831	0,99836	0,99841	0,99846	0,99851	0,99856	0,99861
3	0,99865	0,99869	0,99874	0,99878	0,99882	0,99886	0,99889	0,99893	0,99896	0,999
3,1	0,99903	0,99906	0,9991	0,99913	0,99916	0,99918	0,99921	0,99924	0,99926	0,99929
3,2	0,99931	0,99934	0,99936	0,99938	0,9994	0,99942	0,99944	0,99946	0,99948	0,9995
3,3	0,99952	0,99953	0,99955	0,99957	0,99958	0,9996	0,99961	0,99962	0,99964	0,99965
3,4	0,99966	0,99968	0,99969	0,9997	0,99971	0,99972	0,99973	0,99974	0,99975	0,99976
3,5	0,99977	0,99978	0,99978	0,99979	0,9998	0,99981	0,99981	0,99982	0,99983	0,99983
3,6	0,99984	0,99985	0,99985	0,99986	0,99986	0,99987	0,99987	0,99988	0,99988	0,99989
3,7	0,99989	0,9999	0,9999	0,9999	0,99991	0,99991	0,99992	0,99992	0,99992	0,99992
3,8	0,99993	0,99993	0,99993	0,99994	0,99994	0,99994	0,99994	0,99995	0,99995	0,99995
3,9	0,99995	0,99995	0,99996	0,99996	0,99996	0,99996	0,99996	0,99996	0,99997	0,99997
4	0,99997	0,99997	0,99997	0,99997	0,99997	0,99997	0,99998	0,99998	0,99998	0,99998

Y que refleja:

$$P(Z \leq k)\,, \quad k \in \left[0, 4'09\right]$$



6.3.1. Cálculo en la $Z \sim \mathcal{N}(0, 1)$ o normal estándar

Para determinar la probabilidad de que la normal tome un valor menor (o igual) que un valor concreto k, deberemos ir a la tabla al valor del cruce de la fila y columna correspondiente. Teniendo en cuenta que la fila es el valor hasta las décimas y la columna las centésimas

Ejemplos:

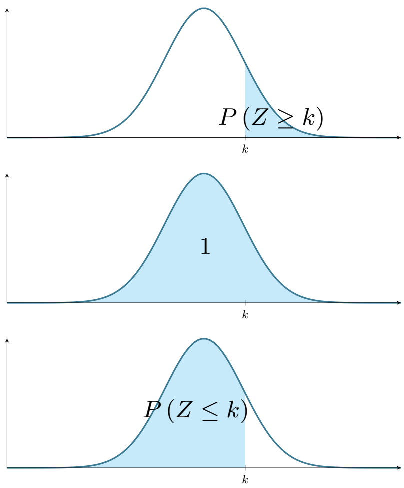
- $P(Z \leq 0) = 0{,}5$. Ya que en este caso $k = 0 = 0{,}00$ la suma del valor de la fila 0 con el valor de la columna 0 me da el valor de k, y la probabilidad asociada es $0{,}5$
- $P(Z \leq 1{,}23) = 0{,}89065$. En este caso la probabilidad asociada a 1.23 se busca en la fila 1.2 y la columna 0.03

z	0	0,01	0,02	0,03	0,04	0,05	0,06
0	0,5	0,50399	0,50798	0,51197	0,51595	0,51994	0,52392
0,1	0,53983	0,5438	0,54776	0,55172	0,55567	0,55962	0,56356
0,2	0,57926	0,58317	0,58706	0,59095	0,59483	0,59871	0,60257
0,3	0,61791	0,62172	0,62552	0,6293	0,63307	0,63683	0,64058
0,4	0,65542	0,6591	0,66276	0,6664	0,67003	0,67364	0,67724
0,5	0,69146	0,69497	0,69847	0,70194	0,7054	0,70884	0,71226
0,6	0,72575	0,72907	0,73237	0,73565	0,73891	0,74215	0,74537
0,7	0,75804	0,76115	0,76424	0,7673	0,77035	0,77337	0,77637
0,8	0,78814	0,79103	0,79389	0,79673	0,79955	0,80234	0,80511
0,9	0,81594	0,81859	0,82121	0,82381	0,82639	0,82894	0,83147
1	0,84134	0,84375	0,84614	0,84849	0,85083	0,85314	0,85543
1,1	0,86433	0,8665	0,86864	0,87076	0,87286	0,87493	0,87698
1,2	0,88493	0,88686	0,88877	0,89065	0,89251	0,89435	0,89617
1,3	0,9032	0,9049	0,90658	0,90824	0,90988	0,91149	0,91309

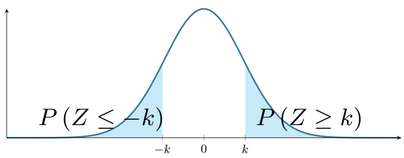
Las probabilidades en conjuntos de valores de la distribución que no se puedan obtener directamente de la tabla se transformarán en operaciones con probabilidades que sí estén en la tabla:

Ejemplos:

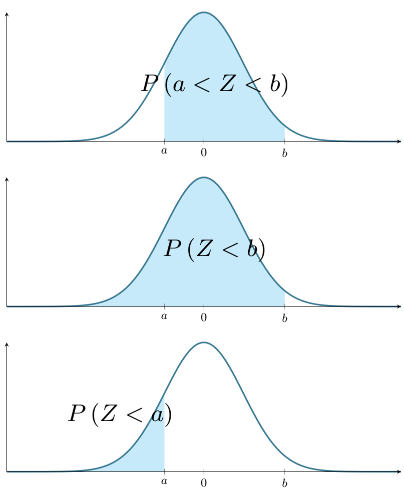
- $P\,(Z \geq 1{,}23) = 1 - P\,(Z \leq 1{,}23) = 1 - 0{,}89065 = 0{,}10935$. Basta fijarse en la relación que hay entre las áreas:



- $P\,(Z \leq -2{,}15) = P\,(Z \geq 2{,}15) = 1 - P\,(Z \leq 2{,}15) = 1 - 0{,}98422 = 0{,}01578$. Basta fijarse en la relación que hay entre las áreas:

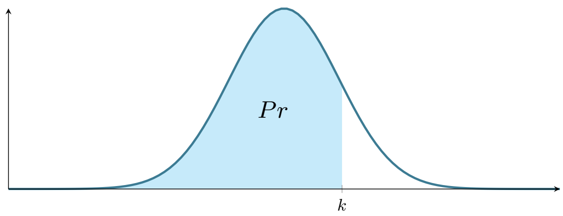


- $P\,(-1{,}3 < Z < 3{,}1) = P\,(Z < 3{,}1) - P\,(Z < -1{,}3) = P\,(Z < 3{,}1) - [1 - P\,(Z < 1{,}3)] = 0{,}99903 - 1 + 0{,}9032 = 0{,}90223$. Basta fijarse en la relación que hay entre las áreas:



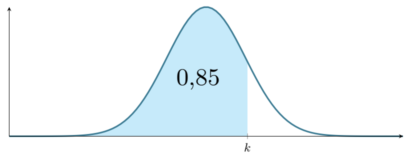
Cálculo del valor de la variable a partir de la probabilidad: El uso de la tabla normal nos permite realizar el proceso inverso. Es decir, fijada una probabilidad Pr, encontrar el valor de la variable k que cumpla:

$$P(Z <= k) = Pr$$



Ejemplo: Dada $Z \sim \mathcal{N}(0,\,1)$, calcula el valor de la variable sabiendo que la probabilidad de que tome un valor menor que ese es de un 85 %.

$$P(Z <= k) = 0{,}85$$



Vamos a la tabla y buscamos los dos valores seguidos de la tabla entre los que se quede el $0{,}85$ y encontramos:

$$0{,}84849 < 0{,}85 < 0{,}85083$$

Como queda más cerca del $0{,}85083$, me quedo con la celda correspondiente y veo que se corresponde con la fila 1 y columna $0{,}04$. Concluimos finalmente que el valor de la variable Z buscado es $k = 1{,}04$.

6.3.2. Cálculo en el resto de normales mediante tipificación

Para manejar $X \sim \mathcal{N}(\mu,\,\sigma)$ reduciremos los cálculos a cálculos en la $Z \sim \mathcal{N}(0,\,1)$ a partir de la siguiente propiedad:

$$Si\ X \sim \mathcal{N}(\mu,\,\sigma) \Rightarrow X^{'} = \frac{X-\mu}{\sigma} \sim \mathcal{N}(0,\,1)$$

Al proceso de transformar la variable anterior a una $Z \sim \mathcal{N}(0,\,1)$ se denomina tipificarla.

Ejemplos:

- $X \sim \mathcal{N}(1,\,2) \rightarrow P(X <= 2{,}32) = P(X' <= \frac{2{,}32-1}{2}) = P(Z <= 0{,}66) = 0{,}74537$
- $X \sim \mathcal{N}(5,\,3) \rightarrow P(X <= 3{,}59) = P(X' <= \frac{3{,}59-5}{3}) = P(Z <= -0{,}47) = 1 - P(Z <= 0{,}47) = 1 - 0{,}68082 = 0{,}31918$