# Distribuciones continuas. Distribución normal.

## Función de densidad.

Cuando X es una variable aleatoria continua, existe una función $f_x$, denominada función de densidad asociada a la variable aleatoria X, que cumple

a) $f_X(x) \geq 0 \ \forall x \in \mathbb{R}$

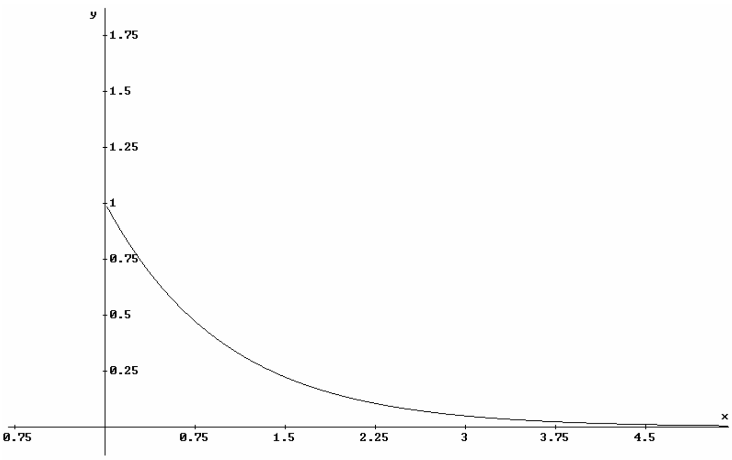
b) $f_X$ admite un número finito de discontinuidades en cada intervalo

c) Si $P_X$ es la función de probabilidad de la v. a. X, se cumplirá

$$P_X(X < x) = \int_{-\infty}^{x} f_X(t).\, dt$$

*# Ejemplo.- Diremos que X es una distribución exponencial de parámetro* $\theta = 1$ *, y lo representamos por* $\exp(1)$ *, si su función de densidad es de la forma*

$$f_X(x) = \begin{Bmatrix} 0 & \text{si } x \leq 0 \\ e^{-x} & \text{si } x > 0 \end{Bmatrix}$$




*Que claramente se comprueba que verifica la Positividad, por ser f una función positiva o nula. Además, también se cumple la normalización, ya que*

$$\int_{-\infty}^{+\infty} f_X(x).\, dx = \int_{0}^{+\infty} e^{-x}.\, dx = 1$$

*Así por ejemplo*

$$P_X(1 \leq X \leq 2) = \int_{1}^{2} e^{-x}.\, dx = e^{-1} - e^{-2}$$

## Media y varianza de una variable aleatoria discreta

Si X es una variable aleatoria continua de función de densidad $f_X$ , denominamos

**Media o esperanza matemática**: $\mu=\int_{-\infty}^{+\infty} x\,.f_X(x)\,.\,dx$

**Varianza**: $\sigma^2=\int_{-\infty}^{+\infty}(x-\mu)^2.\,f_X(x)\,.\,dx$

La raíz cuadrada de la varianza, la denominamos desviación típica $\sigma$ .

*# Ejemplo.- Hallar la media, varianza y desviación típica de una variable aleatoria continua que tiene la siguiente función de densidad*

$$f_X(x)=\begin{cases} 0 & \text{si } x<0 \\ \dfrac{1}{3} & \text{si } 0\le x\le 3 \\ 0 & \text{si } x>3 \end{cases}$$

a) Media: $\mu=\int_0^3 x\,.\frac{1}{3}=\frac{9}{6}=1,5$

b) Varianza: $\sigma^2=\int_0^3 (x-1,5)^2.\frac{1}{3}=0,75$

c) Desviación típica: $\sigma=\sqrt{0,75}=0,866$

## Variable aleatoria de distribución normal

La v. a. continua X sigue la distribución Normal de parámetros $\mu$ y $\sigma$ , y se escribe de forma abreviada $X\equiv N(\mu,\sigma)$ si para cada $x\in\mathbb{R}$ , la probabilidad P cumple

$$P(X\le x)=\int_{-\infty}^{x} f(t)\,.\,dt=\int_{-\infty}^{x}\frac{1}{\sigma.\sqrt{2.\pi}}\,.\,e^{-\frac{(t-\mu)^2}{2.\sigma^2}}\,.\,dt$$

Siendo $\mu$ y $\sigma$ números reales, tal que $\sigma>0$ . Y a la función $f(x)$ se le denomina función de densidad.

La función de densidad $f(x)$ cumple

1.- $f(x)\ge 0\ \forall x\in\mathbb{R}$

2.- $\int_{-\infty}^{+\infty} f(x)\,.\,dx=1$

## Media y varianza de $N(\mu,\sigma)$

Por la propia definición de la distribución, se cumple

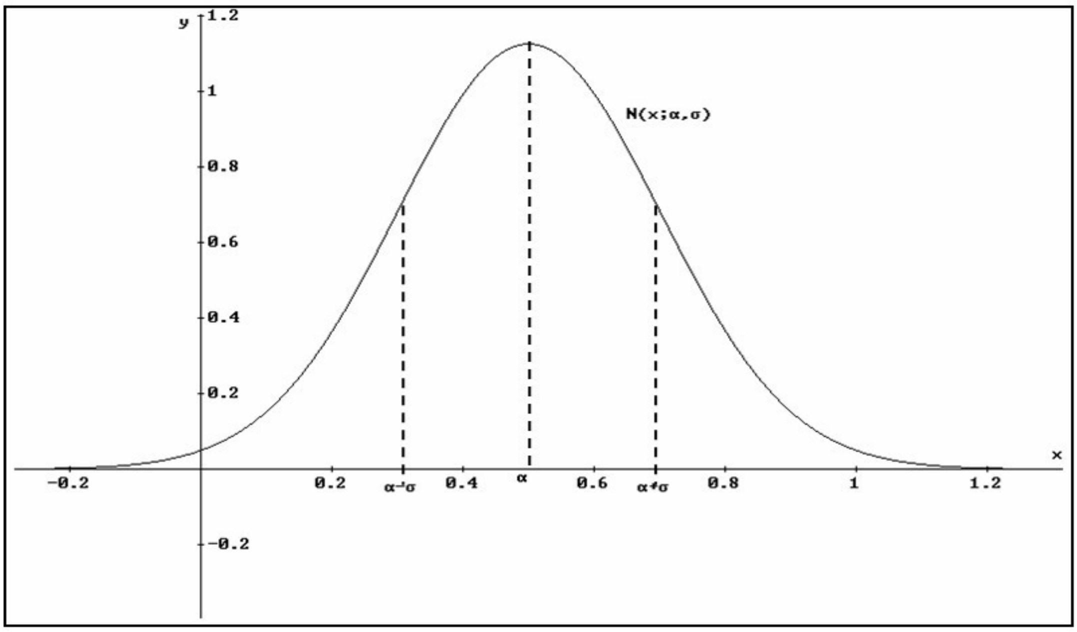
Media $=\mu$ ; Varianza $=\sigma^2$ ; Desviación típica $=\sigma$

## Propiedades de la función de densidad de una distribución $N(\mu, \sigma)$ :

- La función $f(x)$ su valor máximo en $x=\mu$ , es decir f alcanza el máximo en el punto

$$(\mu, f(\mu))=\left(\mu, \frac{1}{\sigma.\sqrt{2.\pi}}\right)$$

- La función $f(x)$ es simétrica respecto del punto $x=\mu$ , es decir $\forall x \in \mathbb{R}$ , se cumple $f(\mu-x)=f(\mu+x)$
- Cuando x tiende a $-\infty$ o $+\infty$ , f(x) tiende a cero
- La función es convexa en $(-\infty, \mu-\sigma)\cup(\mu+\sigma, +\infty)$ y cóncava en $(\mu-\sigma, \mu+\sigma)$
  Es decir, la función es de la forma




## Distribución normal estándar

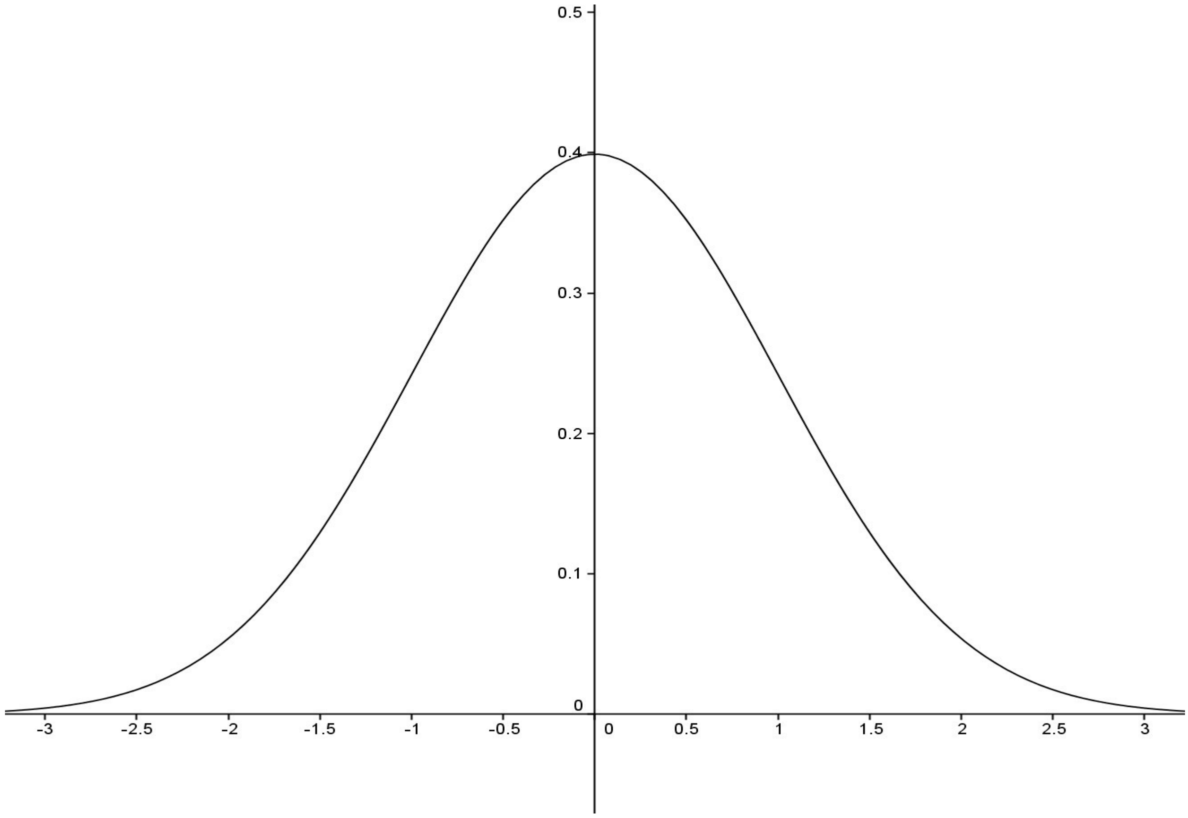
Teniendo en cuenta que dependiendo de que dependiendo de los valores de $\mu$ y de $\sigma$ , la gráfica de la función estará mas desplazada hacia a un lado y más o menos comprimida, y por tanto más o menos alta, siempre podemos utilizar la variable normal tipificada N(0,1), tipificado la variable $X \equiv N(\mu, \sigma)$ a la variable $Y \equiv N(0, 1)$ mediante el cambio o tipificación de la variable

$$Y=\frac{X-\mu}{\sigma}$$

y para cada $y \in \mathbb{R}$, la probabilidad P cumple

$$P(Y \leq y) = \int_{-\infty}^{y} f(t).\, dt = \int_{-\infty}^{y} \frac{1}{\sqrt{2.\pi}} . e^{-\frac{t^2}{2}} .\, dt$$

y cuya función de densidad $f(x) = \frac{1}{\sqrt{2.\pi}} . e^{-\frac{x^2}{2}}$ la podemos representar gráficamente como




Habitualmente calculamos esta probabilidad de forma aproximada mediante las la tabla tabulada de la distribución N(0,1).

*# Ejemplo.- Si las calificaciones de los 500 alumnos de magisterio de Castilla la Mancha presentados a un examen, se distribuye normalmente con media 6,5 y varianza 4.*

*Para calcular la probabilidad de que un aspirante obtenga más de 8 puntos. Como las calificaciones X de los alumnos siguen una distribución normal* $N(6{,}5\,,\sqrt{4}) = N(6\,'5{,}2)$ .

*Y la variable* $Z = \frac{X - \bar{x}}{\sigma} = \frac{X - 6{,}5}{2}$ *sigue una distribución N(0,1).*

*Tenemos* $X > 8 \Rightarrow Z > \frac{8 - 6{,}5}{2}$ *, luego*

$$P(X > 8) = P\left(\frac{8 - 6{,}5}{2}\right) = (Z > 0{,}75) = 1 - P(Z \leq 0{,}75) = 1 - 0{,}7734 = 0{,}2266$$

*Donde el valor P(Z <= 0,75) los hallamos con una tabla de la distribución normal N(1,0)*

**Tabla N(0,1)** $$P(X \le y)=\int_{-\infty}^{x} f(t).\,dt=\int_{-\infty}^{x} \frac{1}{\sqrt{2.\pi}}.\,e^{-\frac{t^2}{2}}.\,dt$$

| 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 | 0,7703 | 0,7734 | 0,7764 | 0,7794 | 0,7823 | 0,7652 |
| **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 | 0,8930 |
| **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 | 0,9561 | 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 | 0,9934 | 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 | 0,9901 | 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 | 0,9954 | 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 |
| **3,1** | 0,9990 | 0,9991 | 0,9991 | 0,9991 | 0,9992 | 0,9992 | 0,9992 | 0,9992 | 0,9993 | 0,9993 |
| **3,2** | 0,9993 | 0,9993 | 0,9994 | 0,9994 | 0,9994 | 0,9994 | 0,9994 | 0,9995 | 0,9995 | 0,9995 |
| **3,3** | 0,9995 | 0,9995 | 0,9995 | 0,9996 | 0,9996 | 0,9996 | 0,9996 | 0,9996 | 0,9996 | 0,9997 |
| **3,4** | 0,9997 | 0.9997 | 0,9997 | 0,9997 | 0,9997 | 0,9997 | 0,9997 | 0,9997 | 0,9997 | 0,9998 |
| **3,5** | 0,9998 | 0,9998 | 0,9999 | 0,9999 | 0,9999 | 0,9999 | 0,9999 | 0,9999 | 0,9999 | 0,9999 |

## Ejemplos de Manejo de tablas de distribución $N(0,1)$

*# Ejemplos*

- $P(Z \leq 0{,}45) = 0{,}6736$
- $P(Z > 1{,}24) = 1 - P(Z \leq 1{,}24) = 1 - 0{,}8925 = 0{,}1075$
- $P(Z \leq -0{,}72) = P(Z \geq 0{,}72) = 1 - P(Z < 0{,}72) = 1 - P(Z \leq 0{,}72) = 1 - 0{,}7642 = 0{,}2358$
- $P(0{,}5 \leq Z \leq 1{,}76) = P(Z \leq 1{,}76) - P(Z \leq 0{,}5) = 0{,}9608 - 0{,}6915 = 0{,}2693$
- $$\begin{aligned} P(-1{,}76 \leq Z \leq -0{,}5) &= P(Z \leq -0{,}5) - P(Z \leq -1{,}76) = \\ &= P(Z > 0{,}5) - P(Z > 1{,}76) = (1 - P(Z \leq 0{,}5)) - (1 - P(Z \leq 1{,}76)) = \\ &= P(Z \leq 1{,}76) - (Z \leq 0{,}5) = 0{,}9608 - 0{,}6915 = 0{,}2693 \end{aligned}$$
- $$\begin{aligned} P(-0{,}53 \leq Z \leq 2{,}46) &= P(Z \leq 2{,}46) - P(Z \leq -0{,}53) = \\ &= P(Z \leq 2{,}46) - P(Z > 0{,}53) = P(Z \leq 2{,}46) - (1 - P(Z \leq 0{,}53)) = \\ &= 0{,}9931 - (1 - 0{,}7019) = 0{,}9931 - 0{,}2981 = 0{,}695 \end{aligned}$$

## Intervalo de confianza para la media

Si el parámetro $\mu$ de una distribución normal $N(\mu, \sigma)$ es desconocido, y para estimarlo tomamos el estimador $\bar{x}$, al tomar una muestra aleatoria de tamaño n.

Como la variable aleatoria

$$var\, X \equiv N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

Tipificando, resulta que

$$Z = \frac{X - \mu}{\sqrt{\frac{\sigma}{n}}} \equiv N(0,1)$$

y se cumplirá que para la probabilidad $1 - \alpha$ (*coeficiente de confianza*), existirá un valor crítico $z_{\frac{\alpha}{2}}$, tal que

$$\begin{aligned} P\left[-z_{\frac{\alpha}{2}} < Z < z_{\frac{\alpha}{2}}\right] = P\left[-z_{\frac{\alpha}{2}} < \frac{X - \mu}{\frac{\sigma}{\sqrt{n}}} < z_{\frac{\alpha}{2}}\right] &= \\ = P\left[-z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} < \bar{x} - \mu < z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}\right] &= \\ = P\left[\bar{x} - z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}\right] &= 1 - \alpha \end{aligned}$$

Luego, el intervalo de confianza para p, al nivel de confianza $1 - \alpha$ es

$$\left(\bar{x} - z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} ; \bar{x} + z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}\right)$$

Si $\sigma$ es desconocido y $n \geq 30$, podemos tomar la desviación típica muestral $\hat{s} = \frac{1}{n-1} \cdot \sum_{i=1}^{n} (x_i - \bar{x})^2$ en vez de la desviación típica $\sigma$ y obtendremos el intervalo de confianza

$$\left(\bar{x} - z_{\frac{\alpha}{2}} \cdot \frac{\hat{s}}{\sqrt{n}} ; \bar{x} + z_{\frac{\alpha}{2}} \cdot \frac{\hat{s}}{\sqrt{n}}\right)$$

*# Ejemplo.- Se sabe que la desviación típica de las tallas de los alumnos de una Universidad es igual 5 cm. Se desea estimarla talla media de los alumnos, para lo cual se escoge una muestra de 100 estudiantes y se obtiene la media bar x = 172 cm. Para hallar el intervalo de confianza al nivel 0,90, 0,95 y 0,99.*

*Como*

$$\bar{x} \equiv N\left(\mu, \frac{\sigma}{n}\right) = N\left(\mu, \frac{5}{\sqrt{100}}\right) = N(\mu ; 0{,}5)$$

*Luego, los intervalos de confianza para la talla media de los alumnos al nivel de confianza* $1-\alpha$ *será*

$$\left(\bar{x} - z_{\frac{\alpha}{2}} \cdot \frac{5}{\sqrt{100}} ; \bar{x} + z_{\frac{\alpha}{2}} \cdot \frac{5}{\sqrt{100}}\right) = N(\mu ; 0{,}5)$$

*que para 90%, 95% y 99%, serán*

$\alpha = 0{,}05 \quad \Rightarrow \quad (172 - 1{,}64 \,.\, 0{,}5 \,;\, 172 + 1{,}64 \,.\, 0{,}5) = (171{,}18 \,;\, 172{,}82)$

$\alpha = 0{,}025 \quad \Rightarrow \quad (172 - 1{,}96 \,.\, 0{,}5 \,;\, 172 + 1{,}96 \,.\, 0{,}5) = (171{,}02 \,;\, 172{,}98)$

$\alpha = 0{,}005 \quad \Rightarrow \quad (172 - 2{,}58 \,.\, 0{,}5 \,;\, 172 + 2{,}58 \,.\, 0{,}5) = (170{,}71 \,;\, 173{,}29)$

Para hallar los intervalos de confianza, tanto para la proporción como para la media, mirando el la tabla $N(0{,}1)$, para cada valor $1-\alpha$, obtenemos los valores más utilizados

| $1-\alpha$ | $\alpha$ | $\frac{\alpha}{2}$ | $z_{\frac{\alpha}{2}}$ |
|---|---|---|---|
| 0,8 | 0,2 | 0,1 | 1,28 |
| 0,9 | 0,1 | 0,05 | 1,64 |
| 0,95 | 0,05 | 0,03 | 0,01 |
| 0,99 | 0,01 | 0,001 | 2,58 |

## Relación entre distribución Binomial $B(n,p)$ y distribución $N(\mu,\sigma)$

Cuando n es grande en la distribución discreta $B(n,p)$ podemos utilizar de forma aproximada la distribución $N(n.p,\sqrt{n.p.q})$, particularmente es fiable si $n.p\geq 0{,}5$ y $n.q\geq 0{,}5$

*# Ejemplo.- Sea un experimento de Bernoulli (solo dos resultados posibles) con resultados que denotaremos {a,b}, y probabilidades* $p$, $q=1-p$. *Dado un número entero* $n>1$, *consideremos el experimento consistente en la repetición de n experimentos de Bernoulli. Si definimos la variable:*

$x_i=1$ *, cuando ocurre el suceso a en el experimento i-esimo.*

$x_i=0$ *, cuando ocurre el suceso b en el experimento i-esimo,* $i=1,2,\ldots,n$.

*Teniendo en cuenta que:*

$$\mu=p.1+q.0=p \quad .$$

$$\sigma^2=p.(1-p)^2+q.(0-q)^2=p\,q(q+p)=p.q \quad .$$

*Y denominando:*

$$s_n=(x_1+x_2+\ldots+x_n)$$

*Se cumplirá:*

$$\lim_{n\to+\infty} B(n,p)\to N(p,\sqrt{n.p.q}) \quad .$$