CAPÍTULO 5: MODELIZAR PARA DECIDIR

Estadística - Lic. en Nutrición – UCEL
Marta Ruggieri, Julia Fernández, M. Eugenia Tesser

VARIABLES ALEATORIAS

A partir de un ejemplo veremos la definición de Variable Aleatoria.

Se considera el experimento aleatorio de arrojar una moneda tres veces y registrar qué lado sale. El espacio muestral en este caso:

S= {(CCC) ,(CCX), (CXC), (XCC), (CXX), (XCX), (XXC), (XXX)}

Si nos interesa estudiar el número de caras al realizar el experimento, entonces se define a la **variable aleatoria** como: *X: número de caras al lanzar tres monedas.*

VARIABLES ALEATORIAS

En cada uno de los puntos muestra de S encontramos los posibles valores que toma la variable aleatoria:

S= {(CCC) ,(CCX), (CXC), (XCC), (CXX), (XCX), (XXC), (XXX)}

x=3 x=2 x=2 x=2 x=1 x=1 x=1 x=0

Al observar cada punto muestra los posibles valores que toma la variable son: *X*=0,1,2,3.

VARIABLES ALEATORIAS

Variable aleatoria es una función de valores reales definida en un espacio muestral. Podemos pensarla como una regla que asigna un (y solo uno) valor numérico a cada punto del espacio muestral, generado por un experimento aleatorio

CLASIFICACIÓN DE VARIABLES ALEATORIAS

Variable aleatoria discreta

- Puede tomar un número finito o infinito numerable de valores.

Variable aleatoria continua

- Puede tomar cualquier valor de un intervalo o conjunto de intervalos.

DISTRIBUCIÓN DE VARIABLES ALEATORIAS

Asociadas a cada variable aleatoria hay un modelo :

Función de probabilidad	• Variable aleatoria discreta • En este curso veremos el modelo **Binomial**
Función de densidad de probabilidad	• Variable aleatoria continua • En este curso veremos el modelo **Normal**

DISTRIBUCIÓN DE VARIABLE ALEATORIA DISCRETA

Una **distribución de probabilidad** está dada por todos los valores que puede tomar la variable aleatoria discreta y las probabilidades asociadas a cada uno de ellos.

Puede representarse mediante una tabla, un gráfico o una fórmula.

La probabilidad de que la variable aleatoria asume un valor, se simboliza: $P(X=x) = p(x)$

DISTRIBUCIÓN DE VARIABLE ALEATORIA DISCRETA

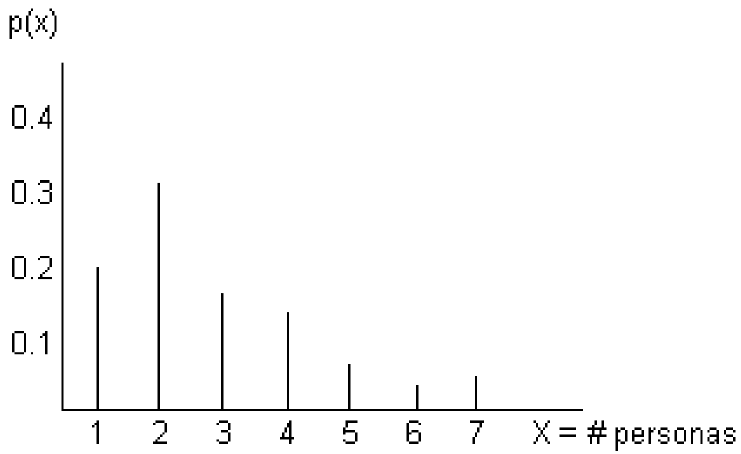
La función de probabilidad de una variable aleatoria discreta X es una función, a menudo indicada con $p(x)= P(X=x)$, que satisface:

- $p(x) \geq 0$ para todo x
- La suma de todos los $p(x)$ es 1.

DISTRIBUCIÓN DE VARIABLE ALEATORIA DISCRETA

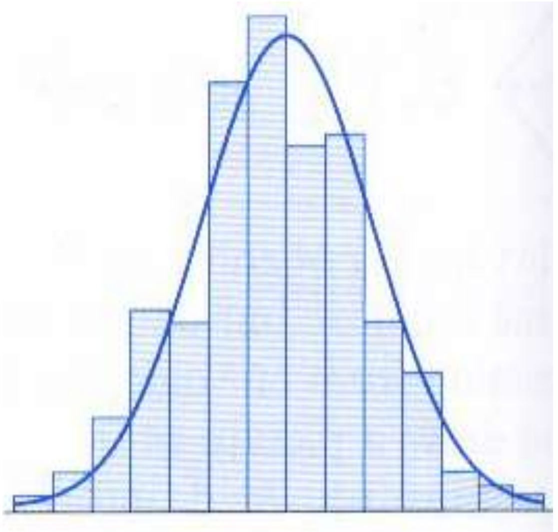
- Ejemplo: Sea X el número de personas internadas por coronavirus en los distintos sanatorios.

X=x	1	2	3	4	5	6	7
P(X=x)	0,20	0,32	0,18	0,15	0,07	0,03	0,05



DISTRIBUCIÓN DE VARIABLE ALEATORIA CONTINUA

La forma de la distribución de una variable aleatoria continua la encontramos didibujando a mano alzada una curva suave que elimine las irregularidades que pueden aparecer en un diagrama de tallos y hojas o un histograma.



En esta figura la curva suavizada es simétrica y campanular, aún cuando los datos son sólo aproximadamente simétricos.

DISTRIBUCIÓN DE VARIABLE ALEATORIA CONTINUA

En un histograma el área de cada rectángulo es proporcional a la frecuencia de cada clase. La curva también proporciona una imagen visual de la proporción en ese área.

Si uno puede obtener la expresión matemática de la curva suavizada tendremos un resumen compacto, simple y razonable de la distribución de los datos.

DISTRIBUCIÓN DE VARIABLE ALEATORIA CONTINUA

Una variable aleatoria continua toma valores en algún intervalo o unión de intervalos. Entonces, no se puede asignar una probabilidad a cada valor posible ya que los valores posibles de un intervalo no se pueden contar. Es decir, la probabilidad asociada a un valor particular es igual a 0.

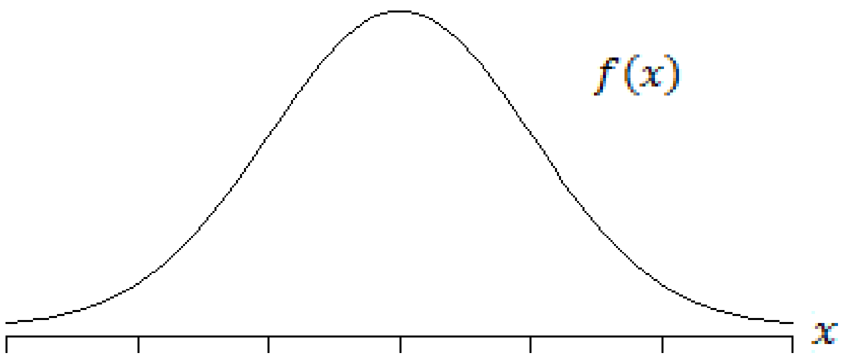
Se asignan probabilidades a intervalos de resultados y se representan esas probabilidades como áreas debajo de la curva, llamada **curva de densidad de probabilidad.**

DISTRIBUCIÓN DE VARIABLE ALEATORIA CONTINUA

El comportamiento de una variable aleatoria continua X viene dado por una función, identificada con **f(x)**, llamada **función de densidad** tal que P(X tome valores en un subconjunto A) = el área bajo la función f(x) sobre el conjunto A.

La función f(x) debe cumplir:

- $f(x) \geq 0$ para todo x
- el área total bajo $f(x)$ es 1.



MANOS A LA OBRA N° 1

VARIABLES ALEATORIAS

Se seleccionan 50 mujeres al azar y se registra el número de frutas que consumen en un día en particular.

a)¿Cuál es el experimento aleatorio?

b) Definir la variable aleatoria en estudio y su clasificación.

c) Se registraron los valores posibles del consumo de frutas en un día y sus respectivas probabilidades.

X=x	0	1	2	3	4	5	6
P(X=x)	0,15	0,22	0,10	0,15	0,35	0	

Completar la tabla.

d)¿Cuál es la probabilidad de seleccionar una mujer al azar y que no haya consumido frutas?

MANOS A LA OBRA N° 1

VARIABLES ALEATORIAS

Se está estudiando la duración (minutos) de cierto medicamento que demora en calmar el dolor de cabeza.

a) Definir la variable en estudio y su clasificación.

b) ¿La función de probabilidad es adecuada para representar la distribución de esta variable?

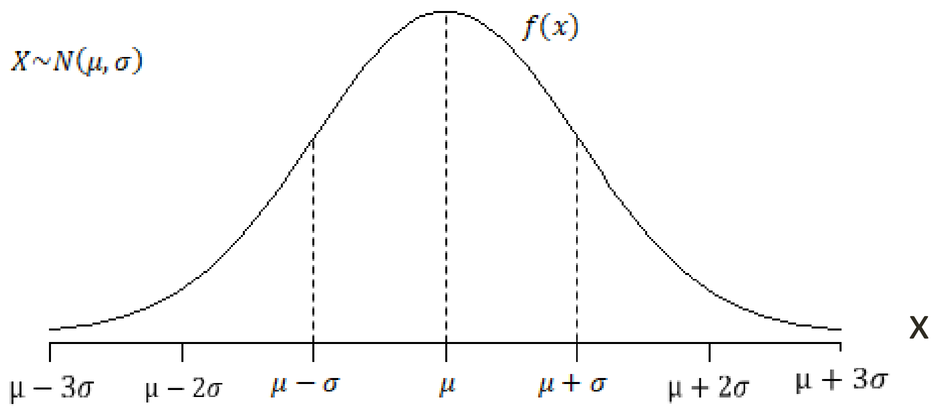
c) Si sabemos que el tiempo en calmar el dolor de cabeza varía de 15 minutos a 90 minutos. ¿Es posible calcular la probabilidad de que una mujer seleccionada al azar se le pase el dolor de cabeza exactamente a los 30 minutos? ¿Y la probabilidad de que el tiempo en calmar el dolor no sea más de 30 minutos pero menos de 60 minutos?

DISTRIBUCIÓN NORMAL

- La curva normal tiene *forma de campana* con un solo máximo justo en el centro de la distribución.
- La media, mediana y moda de la distribución son iguales.
- La mitad del área bajo la curva está a la derecha del pico o máximo, y la otra mitad está a la izquierda.
- El área total bajo la curva es 1.
- Los puntos de inflexión de la curva se dan para $\mu-\sigma$ y $\mu+\sigma$.

DISTRIBUCIÓN NORMAL

- Las probabilidades vendrán dadas por áreas debajo de la curva y el eje x. previa transformación de la variable x en z (variable estandarizada) La Probabilidad en un punto es 0
- La distribución normal es simétrica respecto a su media.
- La distribución normal es asintótica, es decir, la curva se acerca cada vez más al eje *x* pero en realidad nunca llega a tocarlo.



DISTRIBUCIÓN NORMAL

Esta distribución es muy importante por tres razones:

- Numerosos fenómenos continuos parecen seguir esta distribución o pueden aproximarse a ella.
- Proporciona la base para la Inferencia Estadística.
- Se puede utilizar para aproximar distribuciones de probabilidad discreta.

DISTRIBUCIÓN NORMAL

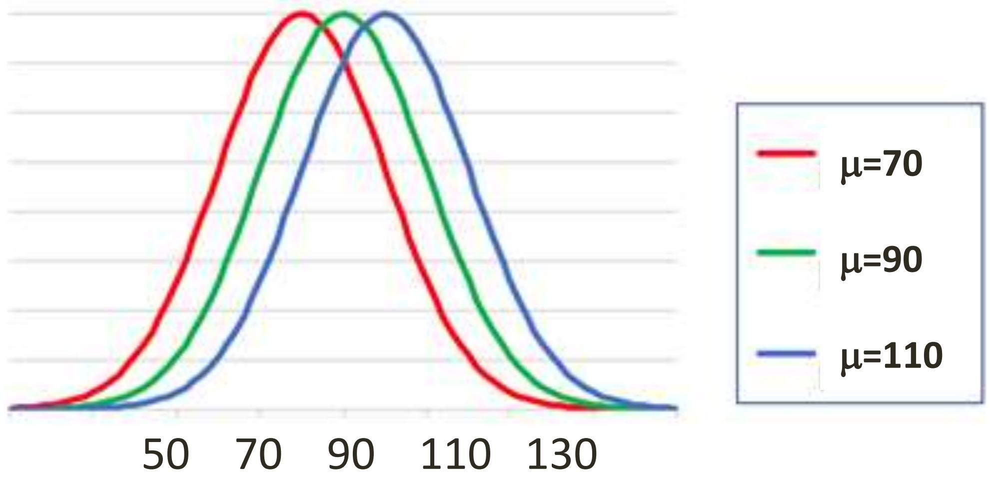
❖ **Notación general**

X es N(μ,σ) significa que la variable o característica X está distribuida normalmente con media μ y desvío estándar σ.

DISTRIBUCIÓN NORMAL

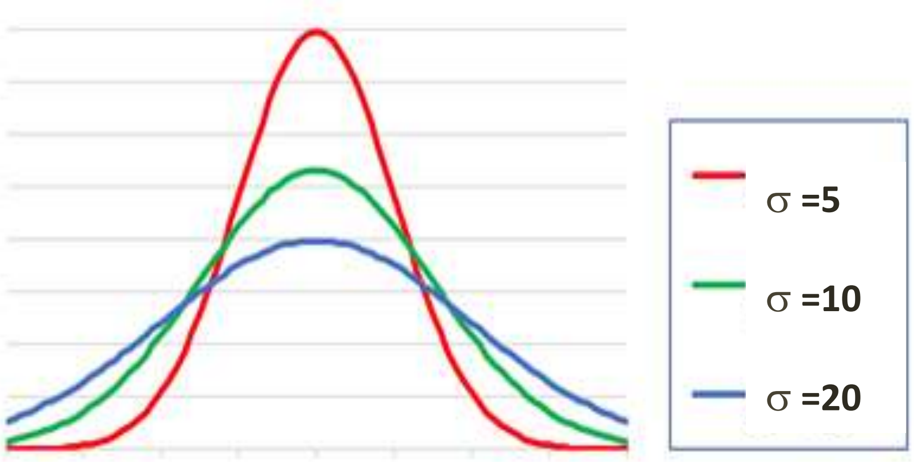
- Ejemplo

El nivel de azúcar en sangre en hombres entre 35 y 50 años sigue una distribución normal:



Podemos observar que si el nivel medio de azúcar en sangre (parámetro μ) aumenta, la función de densidad se traslada a la derecha; en el caso contrario, se traslada a la izquierda.

DISTRIBUCIÓN NORMAL



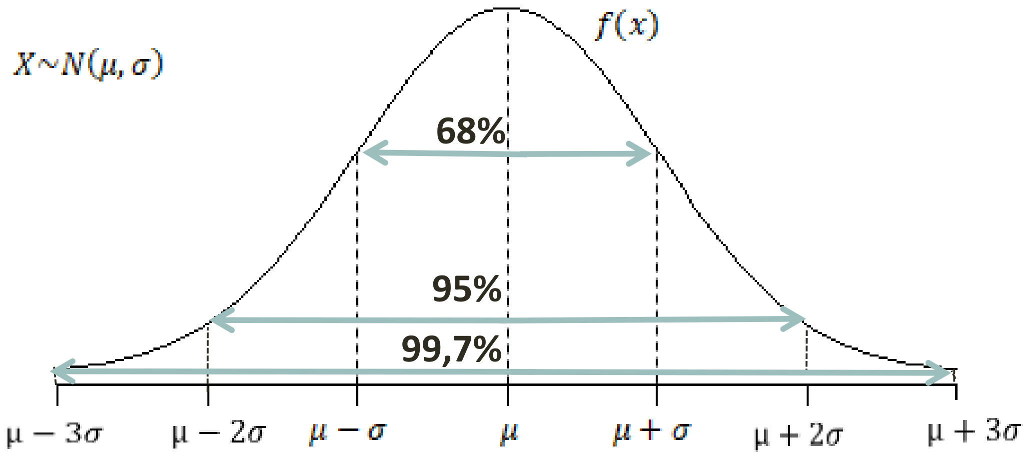
Si el parámetro σ aumenta la curva se “aplana”, los datos se extienden o dispersan a lo largo del eje X; en el caso opuesto, la curva se “apuntala”, los datos se acumulan en el centro de la distribución.

DISTRIBUCIÓN NORMAL

En una distribución Normal, aproximadamente:

- El 68% de las observaciones se encuentra entre $\mu \pm \sigma$. $P(\mu - \sigma \leq X \leq \mu + \sigma) = 0{,}68$
- El 95% de las observaciones se encuentra entre $\mu \pm 2\sigma$. $P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) = 0{,}95$
- El 99,7% de las observaciones se encuentra entre $\mu \pm 3\sigma$. $P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) = 0{,}997$

Regla empírica: 68-95-99.7%



MANOS A LA OBRA N° 2

DISTRIBUCIÓN NORMAL

1) Se supone que el nivel de colesterol para mujeres sanas tiene una distribución normal con μ =160 mg/dl σ=10 mg/dl

a) ¿Entre qué valores de colesterol se encuentra el 95% de los individuos analizados?

b) Además, se está analizando el nivel de colesterol para hombres sanos que también sigue una distribución normal con μ =190 mg/dl σ=18 mg/dl. Si comparamos ambas curvas, ¿Cómo sería la curva de la distribución del colesterol de hombres con respecto a la curva de la distribución del colesterol de mujeres?

MANOS A LA OBRA N° 2

DISTRIBUCIÓN NORMAL

2) Sea X la variable que representa la demora (en minutos) que tiene un paciente hasta que es atendido. Se sabe que X sigue una distribución normal, que la demora promedio es de 15 minutos y el desvío estándar de 5 minutos.

a) Dibujar la curva de densidad para esta distribución.

b) ¿Entre qué valores de demora se encuentra el 99,7% de los pacientes?

MANOS A LA OBRA N° 2

DISTRIBUCIÓN NORMAL

3) Se sabe que el consumo de calorías de adolescentes por día sigue una distribución normal. Cerca del 68% de los adolescentes tienen un consumo de calorías, centrado en la media, entre 2000 calorías y 2200 calorías.

¿Cuál es el valor medio del consumo de calorías por día en adolescentes y el desvío estándar?

DISTRIBUCIÓN NORMAL

Por lo tanto, se puede observar que:

- Un cambio en la media produce un desplazamiento de la curva.
- Un cambio en la desviación estándar altera su forma.

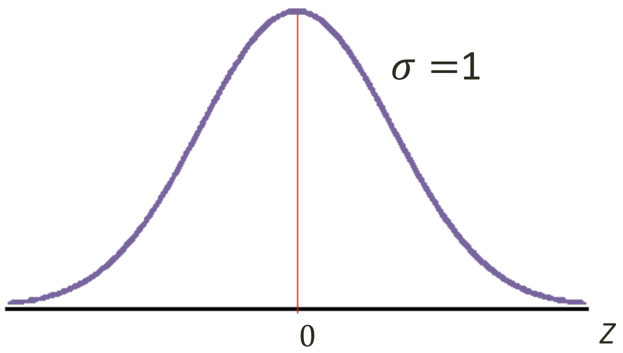
DISTRIBUCIÓN NORMAL ESTÁNDAR

Una distribución normal que tiene media igual a 0 y desviación estándar igual a 1 se denomina **distribución normal estándar**.

Valor *z*: la distancia entre un valor seleccionado, designado como *X*, y la población media μ, dividida entre la desviación estándar de la población σ

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

$Z \sim N(0,1)$



DISTRIBUCIÓN NORMAL ESTÁNDAR

- Ejemplo

Si el nivel de azúcar en sangre para hombres entre 35 años y 50 años sigue una distribución normal con **μ=90 mg/dl y σ=10 mg/dl**

¿Cuál es el valor ***z*** para un nivel de azúcar en sangre de 110 mg/dl? y ¿cuál para uno de 75 mg/dl?

Para X =110, z = (110-90) /10 = 2

z=2 indica que el valor 110 mg/dl está a dos desvíos estándares por encima del nivel medio de azúcar en sangre.

Para X =75, *z* = (75 -90)/10= -1,5

z=-1,5 indica que el valor 75 mg/dl está a 1,5 desvíos estándares por debajo del nivel medio de azúcar en sangre.

DISTRIBUCIÓN NORMAL ESTÁNDAR

La distribución Normal Estándar se encuentra tabulada. Entonces se pueden usar los valores de la misma para encontrar probabilidades bajo cualquier otra distribución normal.

Además, se pueden utilizar distintos programas estadísticos o aplicaciones como por ejemplo Probability Distributions (descargar en el celular).

MANOS A LA OBRA N° 3

<u>DISTRIBUCIÓN NORMAL ESTÁNDAR</u>

Se supone que el nivel de colesterol para mujeres sanas tiene una distribución normal con μ =160 mg/dl σ=10 mg/dl

a) ¿Cuál es el valor ***z*** para un nivel de colesterol de 137 mg/dl?

b) ¿Cuál es el valor ***z*** para un nivel de colesterol de 195 mg/dl?

c) ¿Cuál es el nivel de colesterol para un valor de z=1,2?

CALCULANDO PROPORCIONES NORMALES

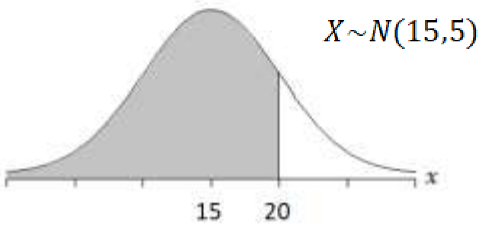
1. Plantea el problema en términos de la variable observada X.
2. Estandariza X para replantear el problema en términos de la variable Normal Estándar Z. Sitúa el área de interés en la curva de la Normal Estándar.
3. Halla el área buscada por debajo de la curva Normal Estándar, utilizando la tabla y considerando que el área total bajo la curva es 1. Además de la tabla se puede utilizar la aplicación Probability Distributions.

CALCULANDO PROPORCIONES NORMALES

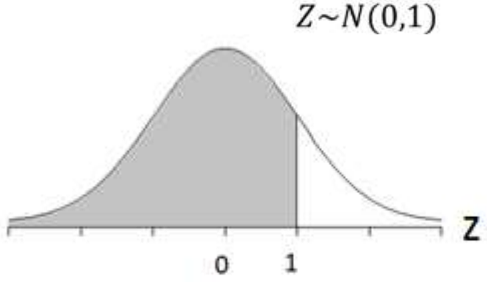
- Ejemplo

Sea X el tiempo de demora (minutos) que tiene un paciente hasta ser atendido. Se asume que X es N(**15 min,5 min**). Interesa conocer ¿cuál es el porcentaje de pacientes que tienen una demora menor o igual 20 minutos?

1. $P(X \leq 20) = P(X < 20)$



2. $Z = \dfrac{X-\mu}{\sigma} = \dfrac{20-15}{5} = 1$



3. P(X<20)=P(Z<1)=0,8413*

*Este valor se halló en la Tabla de la Normal Estándar

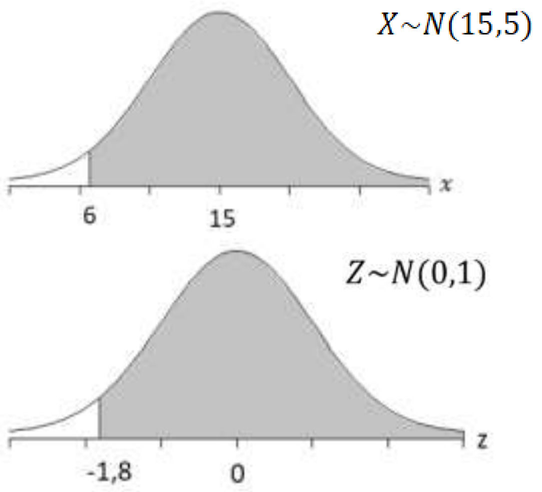
TABLA DE LA NORMAL ESTÁNDAR

z	.00	.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

CALCULANDO PROPORCIONES NORMALES

¿Cuál es el porcentaje de que pacientes que tienen una demora mayor a 6 minutos?

1. P(X>6)

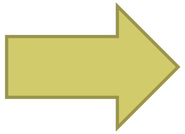


2. $$Z = \frac{X-\mu}{\sigma} = \frac{6-15}{5} = -1{,}8$$

3. P(X>6)=P(Z>-1,8)=1-P(Z<-1,8) =1-0,0359=0,9641

*Este valor se halló en la Tabla de la Normal Estándar

Debido a que la distribución es simétrica y el área debajo de la curva es igual a 1



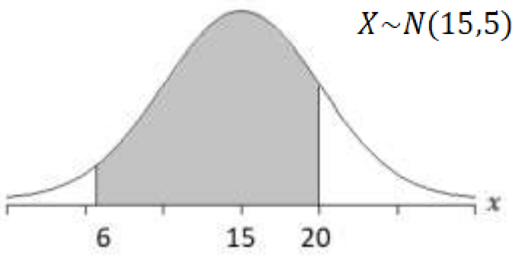
P(Z>-1,8)=1-P(Z<-1,8)

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
−1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
−1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
−1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
−1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
−0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
−0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
−0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
−0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
−0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
−0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
−0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
−0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
−0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
−0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

CALCULANDO PROPORCIONES NORMALES

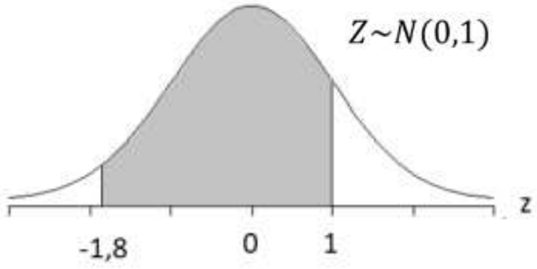
¿Cuál es el porcentaje de pacientes que tienen demoras mayores a 6 minutos pero no más de 20 minutos?

1. P(6<X<20)



2. $$Z_1 = \frac{X-\mu}{\sigma} = \frac{6-15}{5} = -1{,}8$$

$$Z_2 = \frac{X-\mu}{\sigma} = \frac{20-15}{5} = 1$$



3. P(6<X<20)=P(-1,8<Z<1)
   =P(Z<1)-P(Z<-1,8)* =
   0,8413-0,0359=0,8054

*Estos valores se hallaron en la Tabla de la Normal Estándar

CALCULANDO PROPORCIONES NORMALES

Para calcular probabilidades asociadas a una variable normal estandarizada con la tabla correspondiente se procede de la siguiente manera:

- Las probabilidades del tipo $P(Z \leq z)$ se buscan directamente de la tabla.
- Las probabilidades del tipo $P(Z \geq z) = 1 - P(Z \leq z)$
- Las probabilidades de la forma: $P(z_1 \leq Z \leq z_2) = P(Z \leq z_2) - P(Z \leq z_1)$

MANOS A LA OBRA N° 4

CALCULANDO PROPORCIONES NORMALES

1) Se supone que el nivel de colesterol para mujeres sanas tiene una distribución normal con μ =160 mg/dl σ=10 mg/dl.

a) ¿Cuál es el porcentaje de mujeres que tenga al menos 140 mg/dl de nivel de colesterol?

b) Defina el área bajo la curva entre 135 mg/dl y 183 mg/ dl. Interpretar en base al problema.

MANOS A LA OBRA N° 4

CALCULANDO PROPORCIONES NORMALES

2) Se sabe que el consumo de calorías de adolescentes por día sigue una distribución normal con μ =2100 calorías σ=75 calorías

a) ¿Cuál es la proporción de adolescentes que consumen a lo sumo 2000 calorías al día?

b) ¿Cuál es el porcentaje de adolescentes que consumen entre 1950 y 2350 calorías al día.

PERCENTILES DE LA DISTRIBUCIÓN NORMAL

Recordemos…

El p-ésimo percentil es el valor tal que el p% de las observaciones es menor o igual que ese valor.

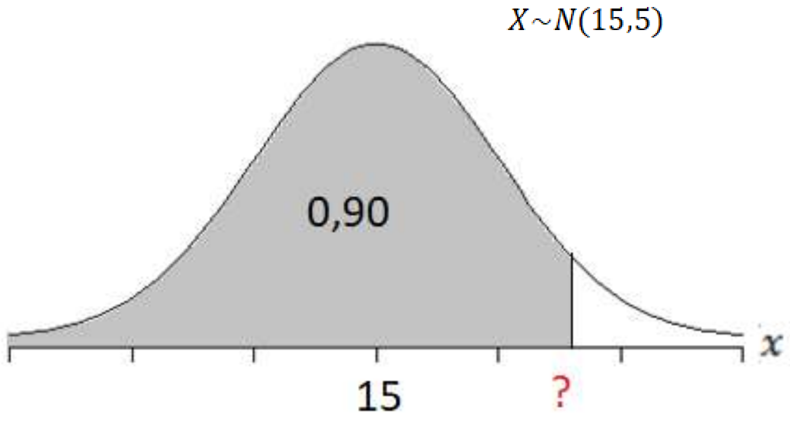
Cada valor de z corresponde a un cierto percentil

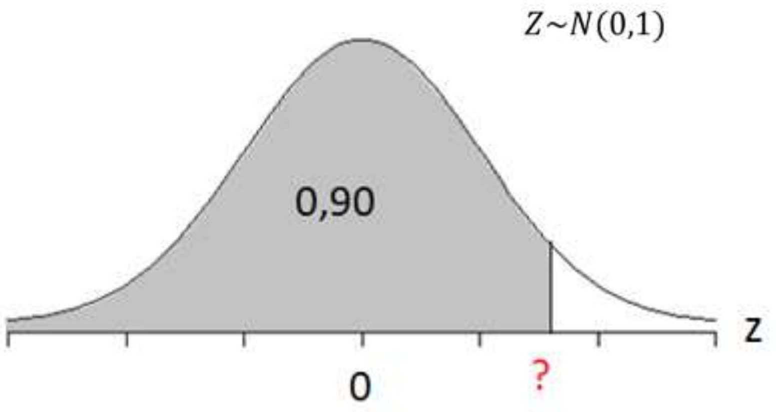
PERCENTILES DE LA DISTRIBUCIÓN NORMAL

- Retomando el ejemplo anterior:

Sea X el tiempo de demora (minutos) que tiene un paciente hasta ser atendido. Se asume que X es N(**15 min,5 min**).

¿Cuántos minutos de demora debe tener el paciente de manera que éste represente el 90% o menos de la distribución?





PERCENTILES DE LA DISTRIBUCIÓN NORMAL

1. P(X<x)=0,90
2. P(Z<z)=0,90 → Busco en la tabla el valor que acumule una probabilidad de 0,90
3. Z=1,282 →

$$z = \frac{X-\mu}{\sigma} \Rightarrow 1{,}282 = \frac{X-15}{5} \Rightarrow 1{,}282\,.\,5 + 15 = X \Rightarrow 21{,}41 = X$$

Un paciente debe tener 21,41 minutos de demora para representar el 90% o menos de la distribución.

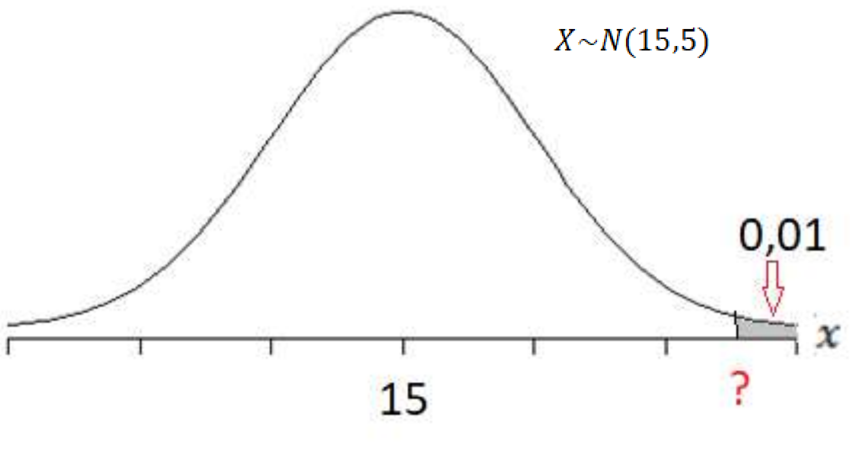
TABLA DE LA NORMAL ESTÁNDAR

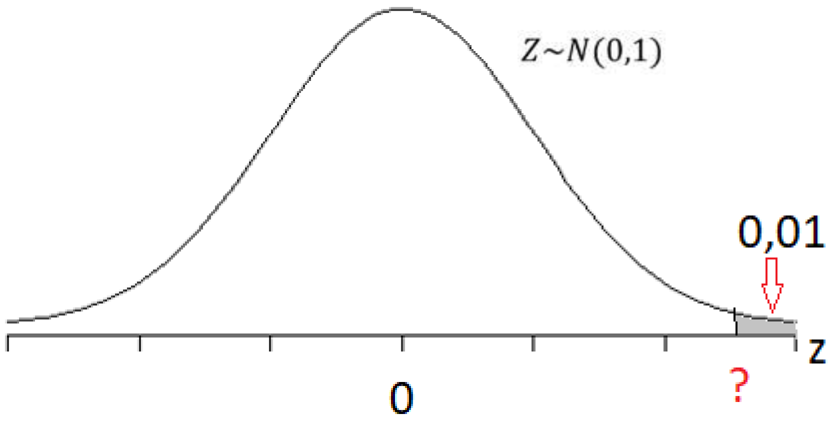
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.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

Area	z
0,00001	-4,265
0,0001	-3,719
0,001	-3,090
0,005	-2,576
0,01	-2,326
0,02	-2,054
0,025	-1,960
0,03	-1,881
0,04	-1,751
0,05	-1,645
0,06	-1,555
0,07	-1,476
0,08	-1,405
0,09	-1,341
0,10	-1,282
0,15	-1,036
0,20	-0,842
0,25	-0,674
0,30	-0,524
0,35	-0,385
0,40	-0,253
0,45	-0,126
0,50	0,000
0,55	0,126
0,60	0,253
0,65	0,385
0,70	0,524
0,75	0,674
0,80	0,842
0,85	1,036
0,90	1,282
0,91	1,341
0,92	1,405
0,93	1,476
0,94	1,555
0,95	1,645
0,96	1,751
0,97	1,881
0,975	1,960
0,98	2,054
0, 99	2,326
0,995	2,576
0,999	3,090
0,9999	3,719
0,99999	4,265

PERCENTILES DE LA DISTRIBUCIÓN NORMAL

¿Cuántos minutos de demora debe tener el paciente de manera que éste represente el 1% superior de la distribución?





PERCENTILES DE LA DISTRIBUCIÓN NORMAL

1. P(X>x)=0,01
2. P(Z>z)=0,01 ⇨ P(Z>z)=1- P(Z<z) ⇨ 0,01=1- P(Z<z) ⇨ P(Z<z)=0,99

   Debido a que el área total bajo la curva es 1 y por la simetría de la distribución

3. Z=2,326 ⇨ $z = \dfrac{X-\mu}{\sigma} \Rightarrow 2{,}326 = \dfrac{X-15}{5} \Rightarrow 2{,}326\,.\,5 + 15 = X \Rightarrow 26{,}63 = X$

Un paciente debe tener 26,63 minutos de demora para representar el 1% superior de la distribución.

MANOS A LA OBRA N° 5

CALCULANDO PROPORCIONES NORMALES

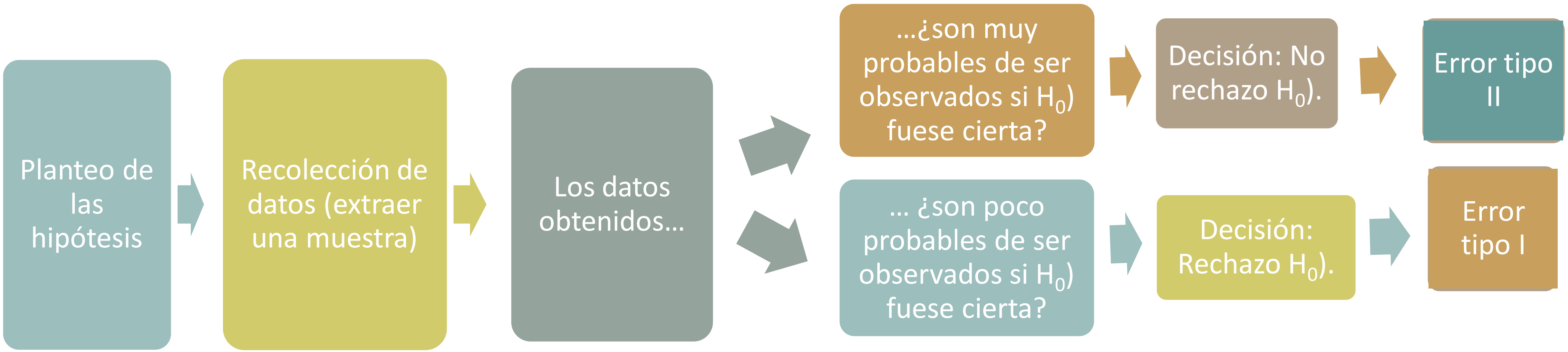
1) Se sabe que el consumo de calorías de adolescentes por día sigue una distribución normal con μ =2100 calorías σ=75 calorías.

a) ¿Cuál es la caloría que representa el 5% superior de la distribución?

b) ¿Cuáles son los valores de calorías que representan el 2,5% inferior de la distribución y el 2,5% superior de la distribución?

PRUEBA DE HIPÓTESIS

Para cualquier situación en la que se contrastan dos hipótesis sobre un parámetro de interés en una población en particular, el proceso de toma de decisiones sigue el siguiente esquema:



PRUEBA DE HIPÓTESIS

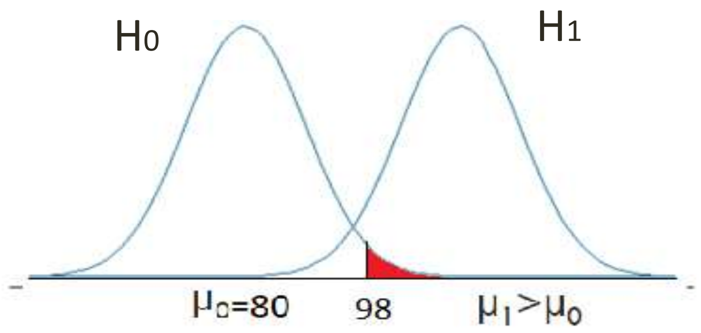
- Se sabe que el nivel medio de azúcar en sangre en adolescentes es 80 mg/dl. Se quiere probar que un grupo que no se alimenta de manera adecuada y saludable tiene mayor nivel medio de azúcar.

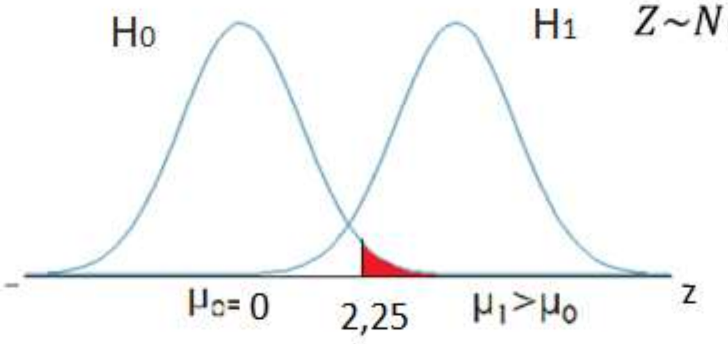
$$H_0)\mu = 80\ mg/dl \qquad\qquad H_1)\mu > 80\ mg/dl$$

EL nivel de azúcar en sangre sigue una distribución normal con desvío estándar igual a 8 mg/dl.

Si se selecciona un adolescente al azar y su nivel de azúcar en sangre es 98 mg/dl ¿Cuál es la probabilidad de que observar este valor o uno más extremos asumiendo que H_0 es cierta?

PRUEBA DE HIPÓTESIS





1. P(X>98)
2. $$z = \frac{X-\mu}{\sigma} = \frac{98-80}{8} = 2{,}25$$
3. P(Z>2,25)=1-P(Z<2,25)=1-0,9878=0,0122

El p-value es igual a 0,0122

PRUEBA DE HIPÓTESIS

- ¿Qué significa p-value=0,0122?
- ¿Es grande o es chico este valor?

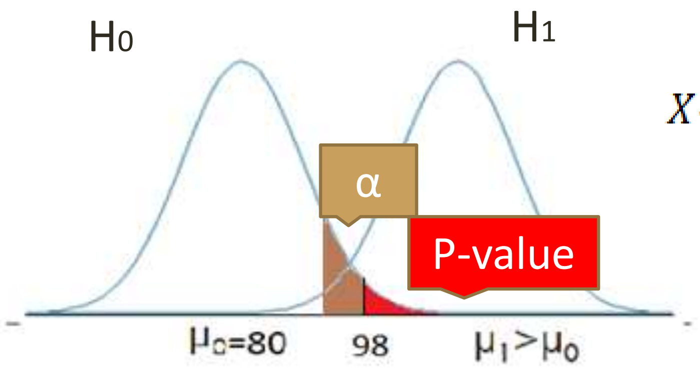
Para poder responder esa pregunta, el investigador fija de antemano un **nivel de significación**, que es la probabilidad de error de tipo 1 (rechazar H_0 siendo cierta) que admite cometer. Los valores más comunes son 0,01 – 0,05 – 0,10.

PRUEBA DE HIPÓTESIS

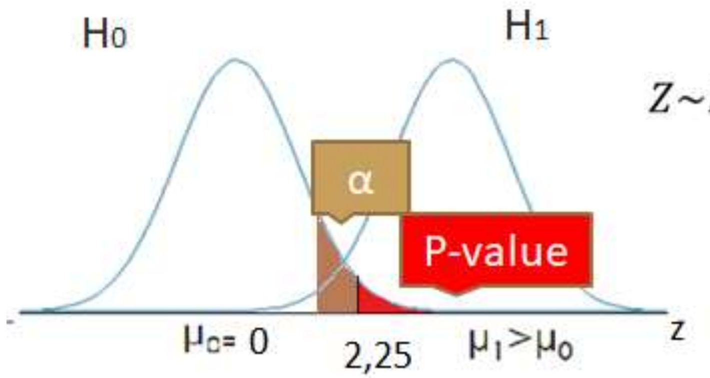
Si el investigador decidiera trabajar con un nivel de significación $\alpha=0,05$ esto implica que asume un 5% de probabilidad de cometer un error de tipo 1 al tomar su decisión, es decir la probabilidad de rechazar H_0 cuando en realidad es cierta es 0,05.

Regla de decisión: Rechazo H_0 si p-value $\leq 0,05$

PRUEBA DE HIPÓTESIS



$X \sim N(80, 8)$



$Z \sim N(0,1)$

Rechazo H_0 si p-value ≤ 0,05

0,0122< 0,05 ⇒ Rechazo H_0

Con un nivel de significación del 5% y en base a la evidencia muestral se puede decir que el nivel medio de azúcar en sangre incrementó

MANOS A LA OBRA N° 6

PRUEBA DE HIPÓTESIS

1) Antes de la implementación de una intervención nutricional en una empresa, la ingesta diaria promedio de frutas y hortalizas de los operarios era 3,5 porciones. Se desea evaluar si luego de la intervención la ingesta diaria promedio de frutas y hortalizas ha incrementado. La variable en estudio es aproximadamente normal con σ=1.

a) Plantear las hipótesis correspondientes.

b) ¿Cuál es la probabilidad de que un operario al azar consuma como máximo 3 porciones?

c) Si un operario seleccionado al azar consume 4 porciones. ¿Cuál es la probabilidad de observar este valor o uno más extremo?

d) Si el investigador considera un nivel de significación del 5% ¿Cuál es la decisión en base al items anterior?

MANOS A LA OBRA N° 6

<u>PRUEBA DE HIPÓTESIS</u>

2) Se sabe que el consumo de calorías de adolescentes por día sigue una distribución normal con μ =2100 calorías σ=75 calorías. Se quiere probar que ciertos alumnos que realizan una dieta estricta en hidratos de carbono consumen menos calorías.

a) Plantear las hipótesis correspondientes.

b) ¿Cuál es el porcentaje de que un adolescente consuma al menos 2200 calorías?

c) Si un adolescente seleccionado al azar consume 1900 calorías. ¿Cuál es la probabilidad de observar este valor o uno más extremo?

d) Si el investigador considera un nivel de significación del 1% ¿Cuál es la decisión en base al items anterior?