Metodología de Investigación

Máster Universitario en Ingeniería Matemática y Computación
Nadia Gámez

Tema 7

UNIVERSIDAD

La Universidad en Internet

Tema 7: La distribución normal

7.1 ¿Cómo estudiar este tema?

Leer Ideas Claves → Sesión Virtual → Lección Magistral

7.2. El modelo de distribución normal

Medidas de Forma

- Nos ayudan a saber la **estructura** de la distribución
- Si es simétrica o asimétrica, puntiaguda o achatada
  - Medidas de **Asimetría**
  - Medidas de **Curtosis**

Medidas de Concentración

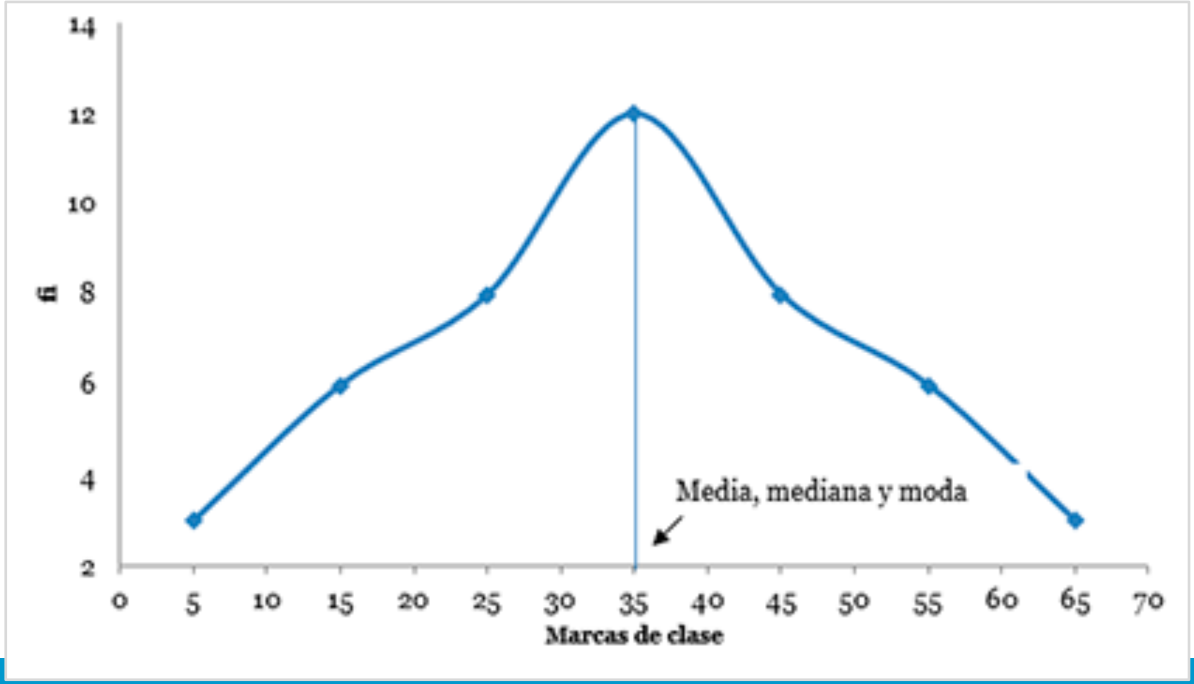
Distribución Normal

7.2. El modelo de distribución normal

Medidas de Forma → **Medidas de Asimetría**

- Una distribución es **simétrica** cuando al doblarla por el eje vertical las ramas de la distribución coinciden

- Distribución **Normal** o **Gausiana**



7.2. El modelo de distribución normal

Medidas de Forma → **Medidas de Asimetría**

- Distribución **simétrica y de variable continua**
  - Media, Mediana o Moda caen en el mismo lugar
  - Dividen la distribución en dos partes de igual frecuencia y valores equidistantes dos a dos

- Distribución **no simétrica → sesgada**
  - Positivamente o negativamente sesgada
  - Moderadamente o significativamente sesgada
  - Unimodal, bimodal o multimodal

7.2. El modelo de distribución normal

Medidas de Forma → **Medidas de Asimetría**

- **Positivamente** sesgada → mayor número de frecuencias a la izquierda y la cola a la derecha
- **Negativamente** sesgada → mayor número de frecuencias están a la derecha y su cola está a la izquierda.

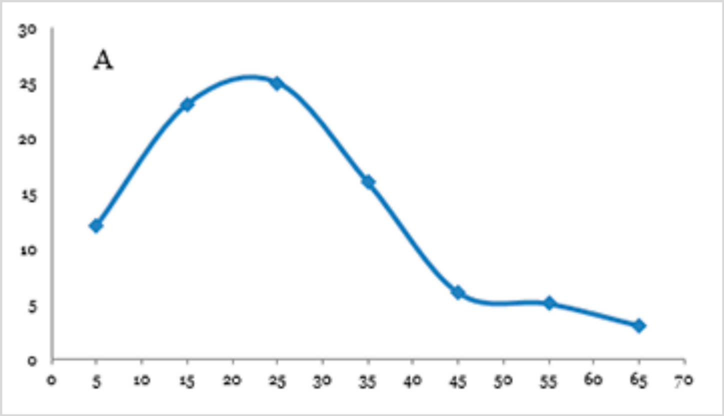
- **Moderadamente** sesgada → poca asimetría
- **Significativamente** sesgada → muy asimétrica

- **Unimodal** → una moda
- **Bimodal** → dos modas
- **Multimodal** → tiene varias modas o picos de máxima concentración

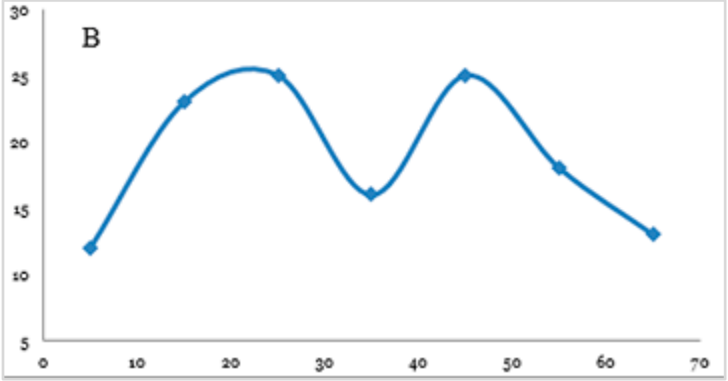
7.2. El modelo de distribución normal

Medidas de Forma → **Medidas de Asimetría**

Distribución moderada, unimodal y positivamente sesgada



Distribución moderada y bimodal

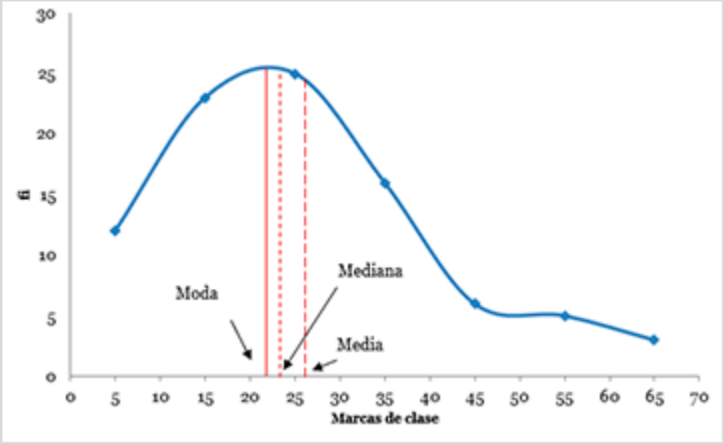


7.2. El modelo de distribución normal

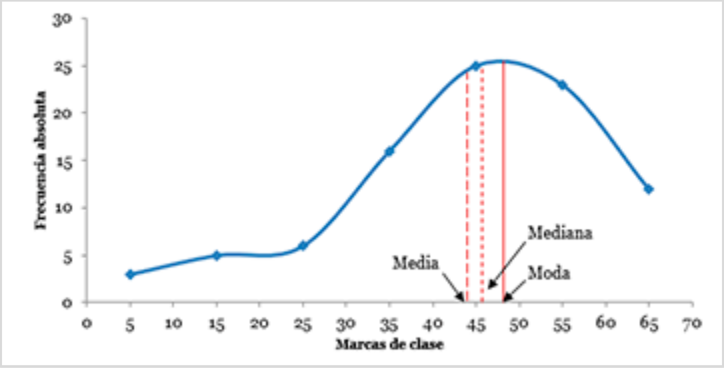
Medidas de Forma → **Medidas de Asimetría**

- Medidas de asimetría → parámetros que miden la asimetría de una distribución respecto a la media, la mediana o la moda.

Distribución positivamente sesgada



Distribución negativamente sesgada



7.2. El modelo de distribución normal

Medidas de Forma → **Medidas de Asimetría**

- El parámetro más empelado para calcular la asimetría es el **coeficiente de Fisher**

$$g_1 = \frac{\dfrac{\sum_{i=1}^{k}(x_i - \bar{x})^3 \cdot f_i}{n-1}}{\left(\sqrt{\dfrac{\sum_{i=1}^{k}(x_i - \bar{x})^2 \cdot f_i}{n-1}}\right)^3}$$

g_1 es el coeficiente de Fisher, $(x_i - \bar{x})$ es la distancia entre los valores y la media, f_i es la frecuencia de cada valor y n el número total de observaciones. El término del denominador es la desviación estándar elevada al cubo.

- Simétrica si es igual a 0
- Positivamente sesgada si mayor que 0
- Negativamente sesgada si menor que 0

Ejemplo 1. Estudiar la asimetría de una distribución

A partir de los datos de la siguiente tabla 1 del tema 2 decidir si hay asimetría y en tal caso el tipo de la misma.

I_i	f_i
4-8	25
9-15	18
16-20	12
21-27	15
28-32	5

Tabla 1. Valores de faltas de asistencia a un curso.

Se calcula la media con las marcas de clase:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{k} x_i \cdot f_i = \frac{6\cdot 25 + 12\cdot 18 + 18\cdot 12 + 24\cdot 15 + 30\cdot 5}{75} = 14{,}56$$

7.2. El modelo de distribución normal

Intervalo	x	f_i	$x-\bar{x}$	$(x-\bar{x})^2$	$f_i \cdot (x-\bar{x})^2$	$(x-\bar{x})^3$	$f_i \cdot (x-\bar{x})^3$
4-8	6	25	-8,56	73,27	1831,84	-627,22	-15680,55
9-15	12	18	-2,56	6,55	117,96	-16,78	-301,99
16-20	18	12	3,44	11,83	142,00	40,71	488,49
21-27	24	15	9,44	89,11	1336,70	841,23	12618,49
28-32	30	5	15,44	238,39	1191,97	3680,80	18403,99
Total		75			4620,48		15528,42

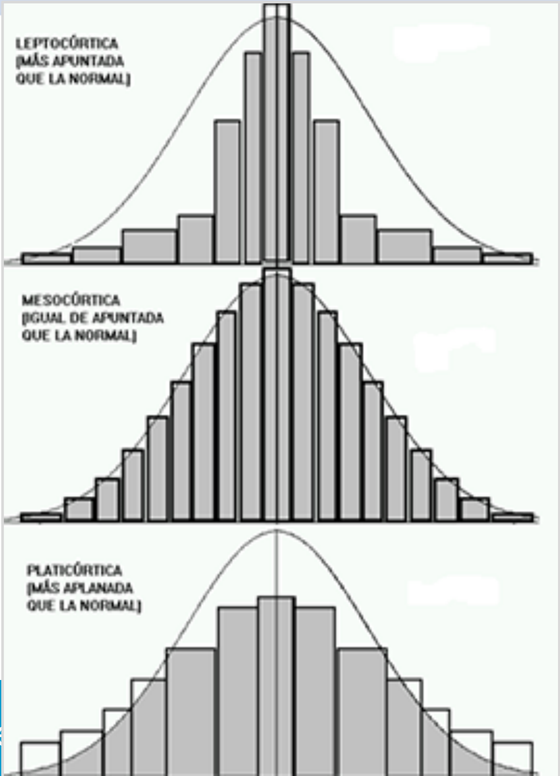
$$g_1 = \frac{\frac{\sum_{i=1}^{k}(x_i - \bar{x})^3 \cdot f_i}{n-1}}{\left(\sqrt{\frac{\sum_{i=1}^{k}(x_i - \bar{x})^2 \cdot f_i}{n-1}}\right)^3} = \frac{\frac{15528{,}42}{74}}{\left(\sqrt{\frac{4620{,}48}{74}}\right)^3} = 0{,}43$$

Puesto que 0,43> 0 la distribución es positivamente sesgada.

7.2. El modelo de distribución normal

Medidas de Forma → **Medidas de Curtosis**

- Información sobre el apuntamiento de una distribución unimodal, campaniforme y moderadamente simétrica
- Si alrededor de la media hay muchos valores, el apuntamiento será elevado.
  - Mesocúrticas→ apuntamiento es el mismo que el de la distribución normal
  - Leptocúrticas → más apuntamiento que la distribución normal
  - Platicúrticas → menos apuntamiento que la distribución normal



7.2. El modelo de distribución normal

Medidas de Forma → **Medidas de Asimetría**

- El grado de apuntamiento se mide mediante el **coeficiente de curtosis**

$$g_2 = \frac{\frac{\sum_{i=1}^{k}(x_i - \bar{x})^4 \cdot f_i}{n-1}}{\left(\sqrt{\frac{\sum_{i=1}^{k}(x_i - \bar{x})^2 \cdot f_i}{n-1}}\right)^4} - 3$$

- Normal o mesocúrtica si igual a 0
- Mucho apuntamiento o leptocúrtica si mayor que 0
- Achatada o platicúrtica si menor que 0

7.2. El modelo de distribución normal

Ejemplo 2. Calcular el apuntamiento de una distribución

A partir de los datos de la siguiente tabla 1 del tema 2 estudiar el apuntamiento de la distribución:

I_i	f_i
4-8	25
9-15	18
16-20	12
21-27	15
28-32	5

Tabla 2. Valores de faltas de asistencia a un curso.

Se calcula la media con las marcas de clase:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{k} x_i \cdot f_i = \frac{6 \cdot 25 + 12 \cdot 18 + 18 \cdot 12 + 24 \cdot 15 + 30 \cdot 5}{75} = 14{,}56$$

7.2. El modelo de distribución normal

Intervalo	x	f_i	$x-\bar{x}$	$(x-\bar{x})^2$	$f_i \cdot (x-\bar{x})^2$	$(x-\bar{x})^4$	$f_i \cdot (x-\bar{x})^4$
4-8	6	25	-8,56	73,27	1831,84	5369,02	134225,51
9-15	12	18	-2,56	6,55	117,96	42,95	773,09
16-20	18	12	3,44	11,83	142,00	140,03	1680,41
21-27	24	15	9,44	89,11	1336,70	7941,23	119118,51
28-32	30	5	15,44	238,39	1191,97	56831,51	284157,54
Total		75			4620,48		539955,06

$$g_2 = \frac{\frac{\sum_{i=1}^{k}(x_i - \bar{x})^4 \cdot f_i}{n-1}}{\left(\sqrt{\frac{\sum_{i=1}^{k}(x_i - \bar{x})^2 \cdot f_i}{n-1}}\right)^4} - 3 = \frac{\frac{539955,06}{74}}{\left(\sqrt{\frac{4620,48}{74}}\right)^4} - 3 = -1,13$$

Puesto que -1,13 < 0 la distribución es platicúrtica o achatada.

7.2. El modelo de distribución normal

Medidas de Concentración

- Dan información sobre el **grado de desigualdad** que tiene una variable en un conjunto.

- Dos tipos de concentración:
  - **Mínima concentración** es similar a una situación de **máxima equidad** entre los valores
  - **Máxima concentración** supone **mínima equidad**

- Para saber el grado de concentración se emplea:
  - **Índice de Gini**, que da valores numéricos
  - **Curva de Lorenz**, que da valores gráficos

7.2. El modelo de distribución normal

Medidas de Concentración → Índice de Gini

- Ordenar los datos

Valor	Individuos	Individuos acumulados	Valor obtenido	Valor acumulado	$p_i = \frac{precep.acum.}{n} = \frac{F_i}{n}$	$q_i = \frac{valor.acum.}{valor\ total} = \frac{Ui}{\sum_{i=1}^{k} x_i \cdot f_i}$
x_1	f_1	$f_1=F_1$	$x_1 \cdot f_1 = V_1$	$V_1=U_1$	p_1	q_1
x_2	f_2	$f_1+f_2=F_2$	$x_2 \cdot f_2 = V_2$	$V_1+V_2 = U_2$	p_2	q_2
...	...	...	...	....	...	...
x_i	f_i	$f_1+f_2+...=F_i$	$x_i \cdot f_i = V_i$	$V_1+V_2+....V_i= U_i$	p_i	q_i
...	...	...	...	....	...	...
x_k	f_k	$f_1+f_2+...f_i+... =F_k$	$x_k \cdot f_k = V_k$	$V_1+V_2+....V_i+ ...V_k=U_k$	p_k	q_k
	n		$\sum_{i=1} x_i \cdot f_i = V$	U		

$$I_G = \frac{\sum_{i=1}^{n-1}(p_i - q_i)}{\sum_{i=1}^{n-1} p_i}$$

7.2. El modelo de distribución normal

Ejemplo 3. Cálculo del Índice de Gini

A partir de los datos de salarios recibidos por empleados en una empresa calcular el índice de concentración de Gini.

Intervalo	f_i
8.000-10.000	35
10.000-15.000	64
15.000-20.000	23
20.000-30.000	3
30.000-50.000	7

Tabla 4. Salarios anuales recibidos por los empleados de una empresa.

7.2. El modelo de distribución normal

x_i	f_i	F_i	$x_i \cdot f_i$	U	$p_i = \frac{F_i}{n}$	$q_i = \frac{Ui}{\sum_{i=1}^{k} x_i \cdot f_i}$	pi-qi
9.000	35	35	315000	315000	0,26	0,17	0,10
12.500	64	99	800000	1115000	0,75	0,59	0,15
17.500	23	122	402500	1517500	0,92	0,81	0,11
25.000	3	125	75000	1592500	0,95	0,85	0,10
40.000	7	132	280000	1872500	-	-	-
Total	132		1872500		2,89		0,46

$$I_G = \frac{\sum_{i=1}^{n-1}(p_i - q_i)}{\sum_{i=1}^{n-1} p_i} = \frac{0,46}{2,89} = 0,16$$

7.2. El modelo de distribución normal

Medidas de Concentración → **Curva de Lorenz**

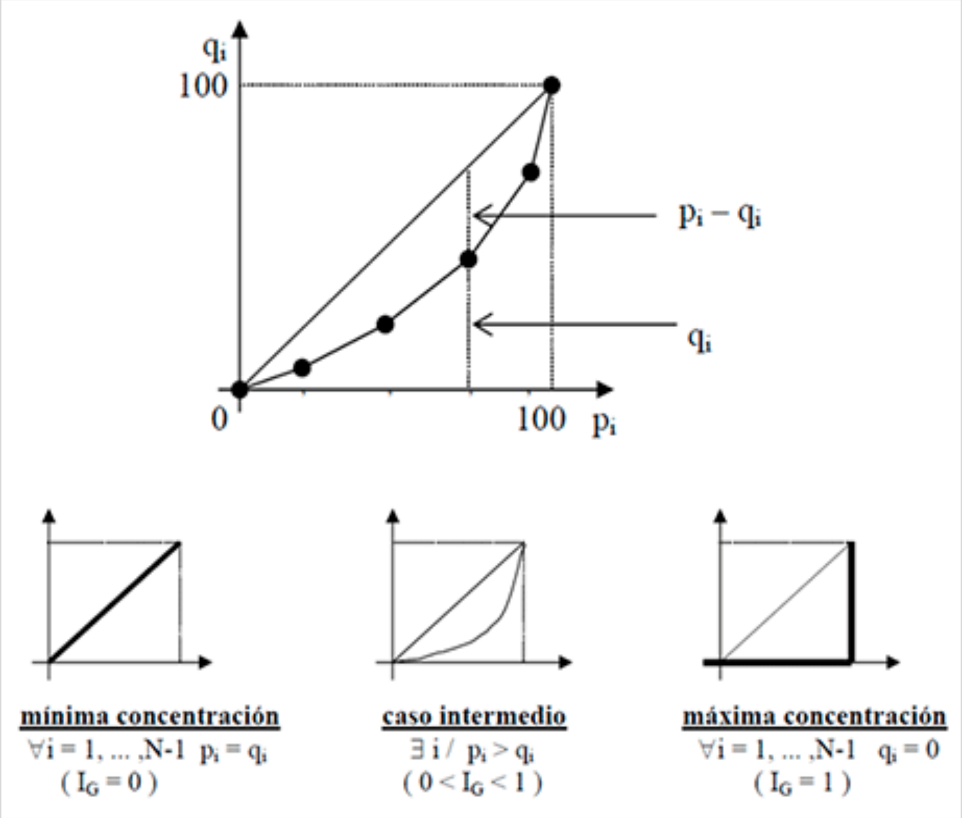
- La curva de Lorenz se obtiene a partir del Índice de Gini
- Permite visualizar la manera en la que se han llevado a cabo los repartos
- Para dibujar la curva, se colocan en el **eje de abscisas los valores de pi y los valores de qi en el eje de ordenadas**
- Los valores de pi y qi oscilan entre 0 y 1, o entre o y 100 si se pone en porcentaje.
- Al unir los puntos se obtiene una curva que es la curva de Lorenz.

$$I_G = \frac{\text{área entre curva de Lorenz y la diagonal}}{\text{área de la diagonal con los ejes}}$$

7.2. El modelo de distribución normal

Medidas de Concentración → Curva de Lorenz

- **Mínima concentración**: recta de equidistribución (valores de pi iguales a los qi)
- **Máxima concentración**: la curva de Lorenz se transforma en los catetos de un triángulo rectángulo (diferencia entre los valores de pi y qi es la mayor posible y la curva está lo más alejada posible de la diagonal)
- **Concentración intermedia**: la curva de Lorenz es una función cuadrática. Cuanto más se aleje la curva de la diagonal más situación de máxima concentración habrá.



7.2. El modelo de distribución normal

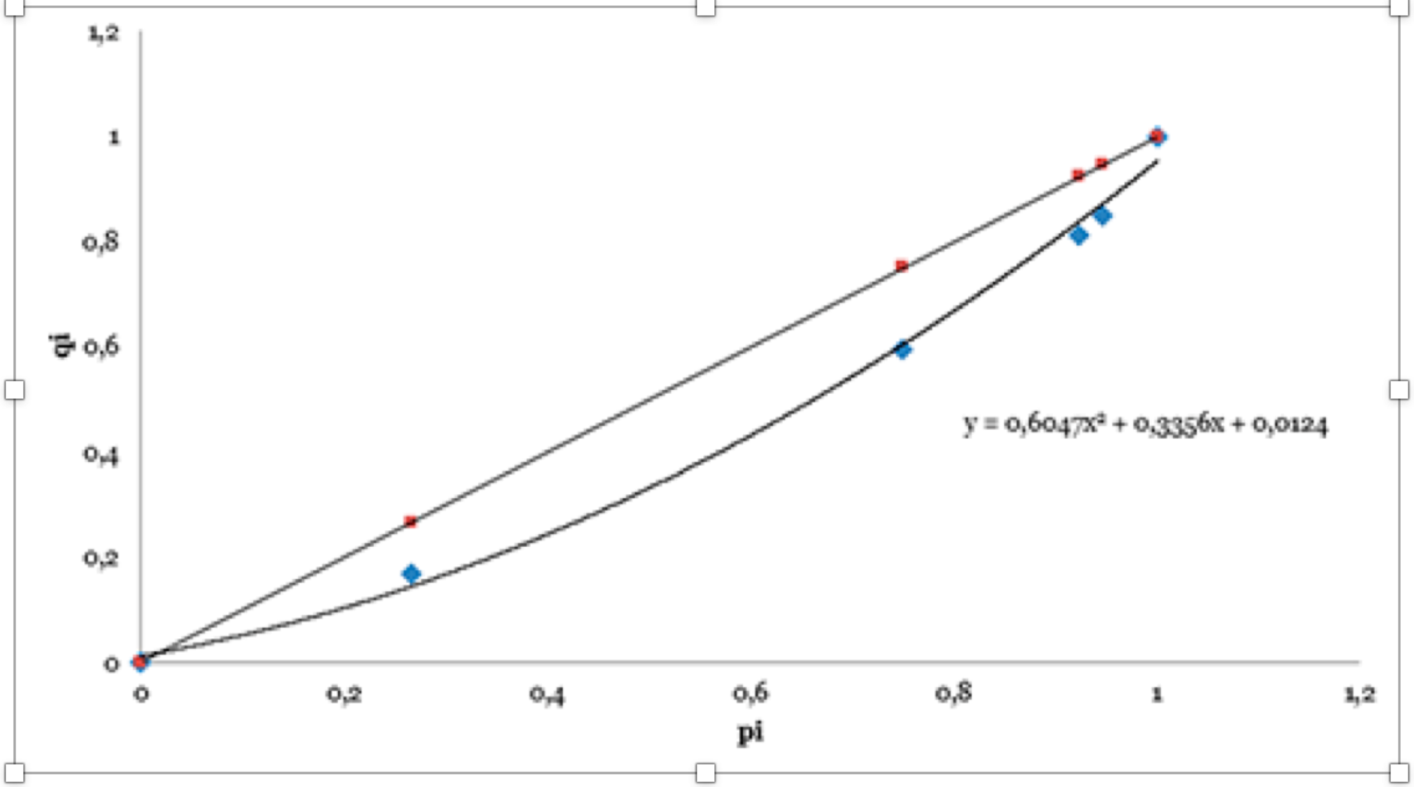


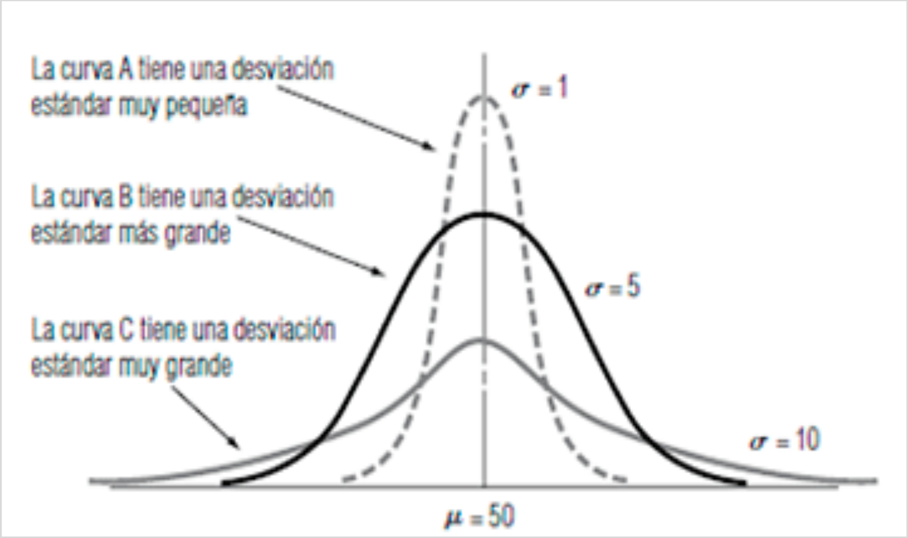
Gráfico 6. Curva de Lorenz para el Ejemplo 5.

La situación es de mínima concentración debido a que el área es pequeña.

7.2. El modelo de distribución normal

La Distribución Normal

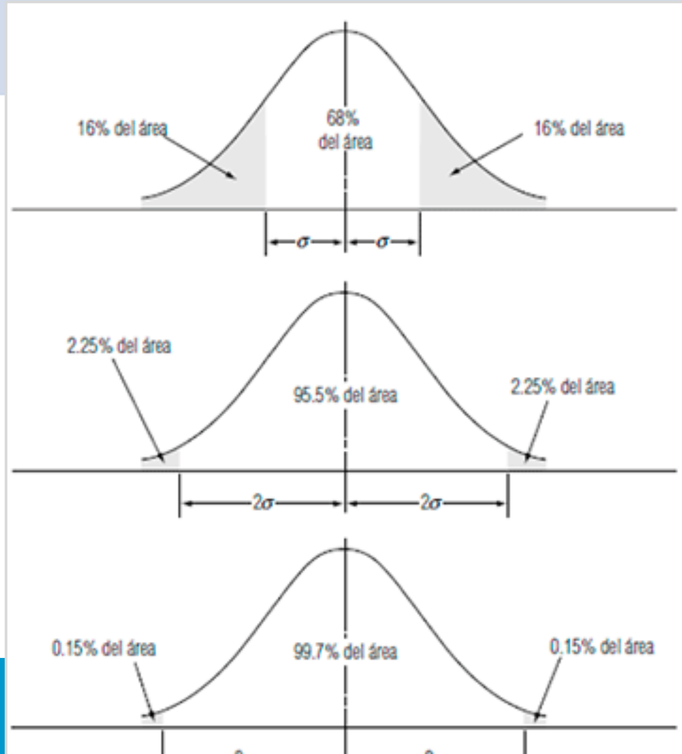
- Muchas situaciones reales se ajustan a este tipo de distribución
- Util en en las distribuciones de **muestreo**
- La **media cae en el punto medio** de la distribución junto con la mediana y la moda y es unimodal
- Las **dos colas simétricas** nunca tocan al eje horizontal de la distribución por lo que la probabilidad se extiende indefinidamente



7.2. El modelo de distribución normal

La Distribución Normal

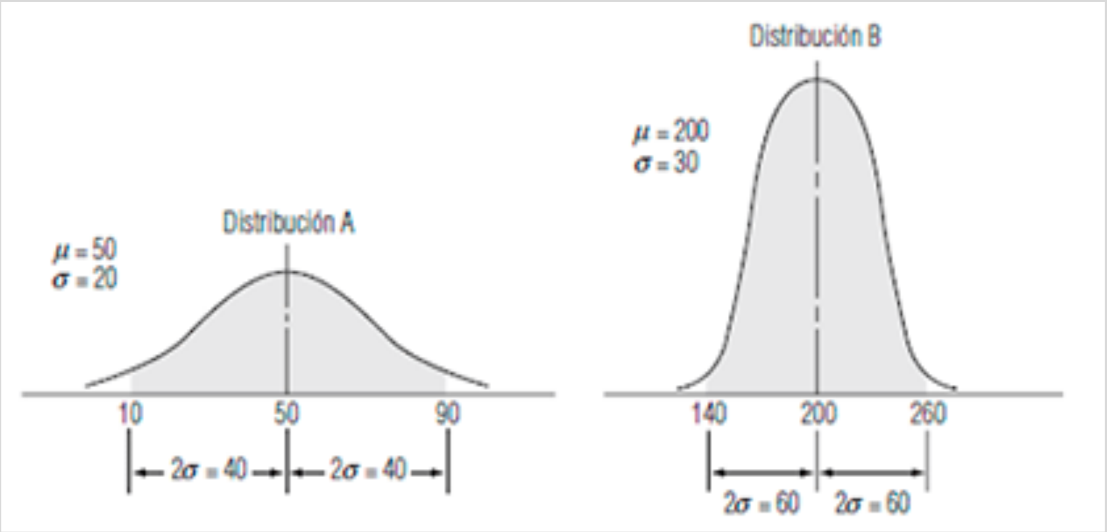
- Independientemente de los valores de las media y la desviación estándar de la población, el **área total bajo la curva siempre es 1** o el 100% y las diferentes áreas que se vaya tomando son probabilidades. Se demuestra que:
- El 68% aprox de los datos caen a ±1 desviación estándar de la media
- El 95% aprox de los datos caen a ±2 desviaciones estándar de la media
- El 99,7% aprox de los datos caen a ±3 desviaciones estándar de la media



7.2. El modelo de distribución normal

La Distribución Normal

- La mayor parte de las situaciones caerán en valores de distancias en desviaciones estándar de la media diferentes
- Usar de tablas que nos permitan saber la probabilidad de los datos
- Dos distribuciones normales que tengan diferente media y diferente desviación estándar pueden tener el mismo área bajo la curva si las áreas sombreadas están a los mismos valores de desviaciones estándar de la media
- Distribución de **probabilidad normal estándar** para encontrar las áreas bajo la curva de cualquier curva normal



7.3. La estandarización Z

Estandarización mediante puntuaciones Z

- Lo primero es introducir la **estandarización** mediante un parámentro que se denota por **Z**
- Este parámetro se emplea como sustituyente al número de desviaciones estándar respecto a la media, con el objetivo de hacer una estandarización de tales valores como números **enteros de 0 a 3** en vez de tener que calcular para cada caso en particual el valor de las desviaciones estándar
- Para calcular Z se emplea la siguiente fórmula:

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

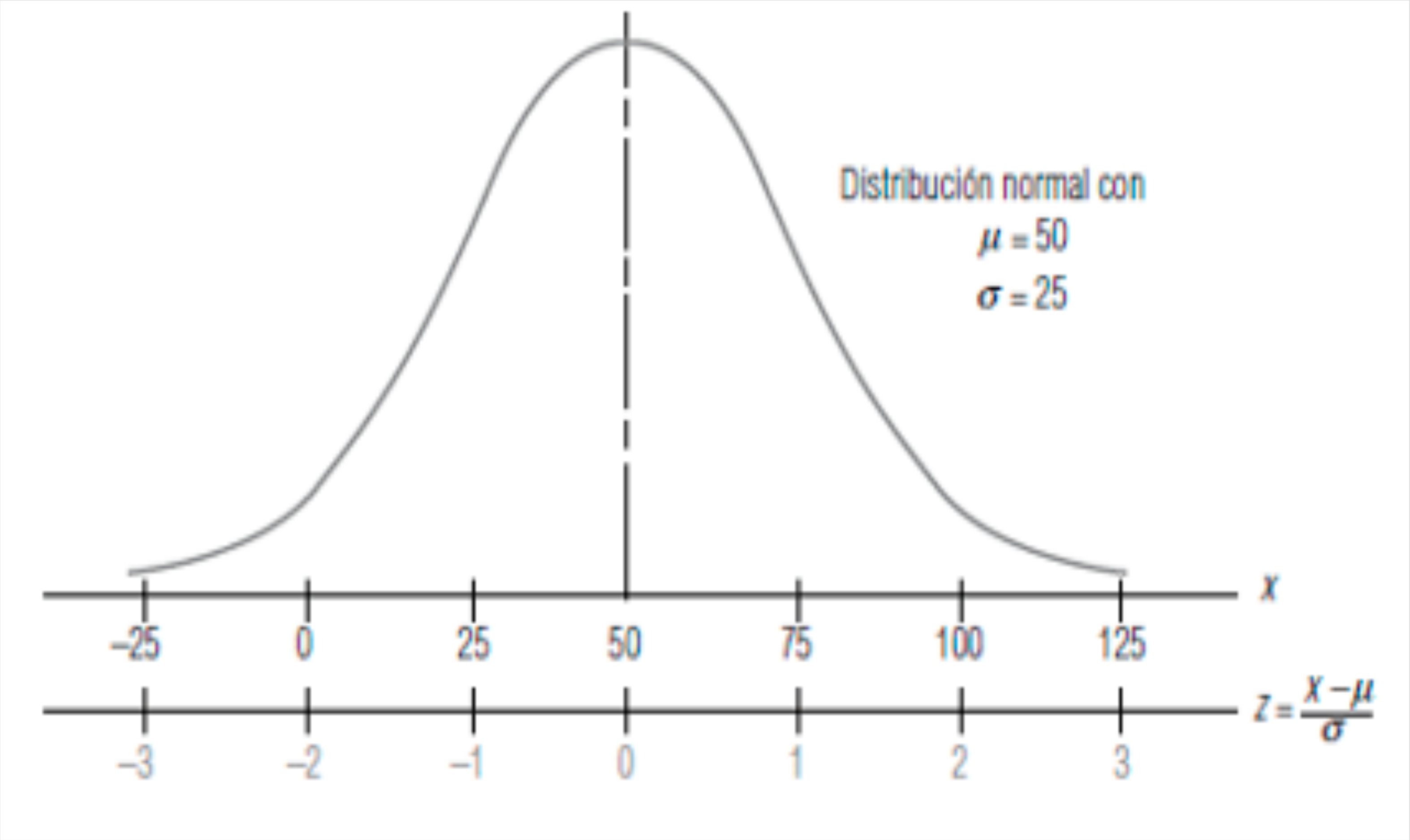
x es la variable aleatoria en cuestión, μ y σ son la media y la desviacion estándar de la distribución y Z es el número de desviaciones estándar a las que la variable de estudio se halla de la media.

7.3. La estandarización Z

Estandarización mediante puntuaciones Z

- Lo
pa
- Es
de
ha
en
pa
- Pa

en



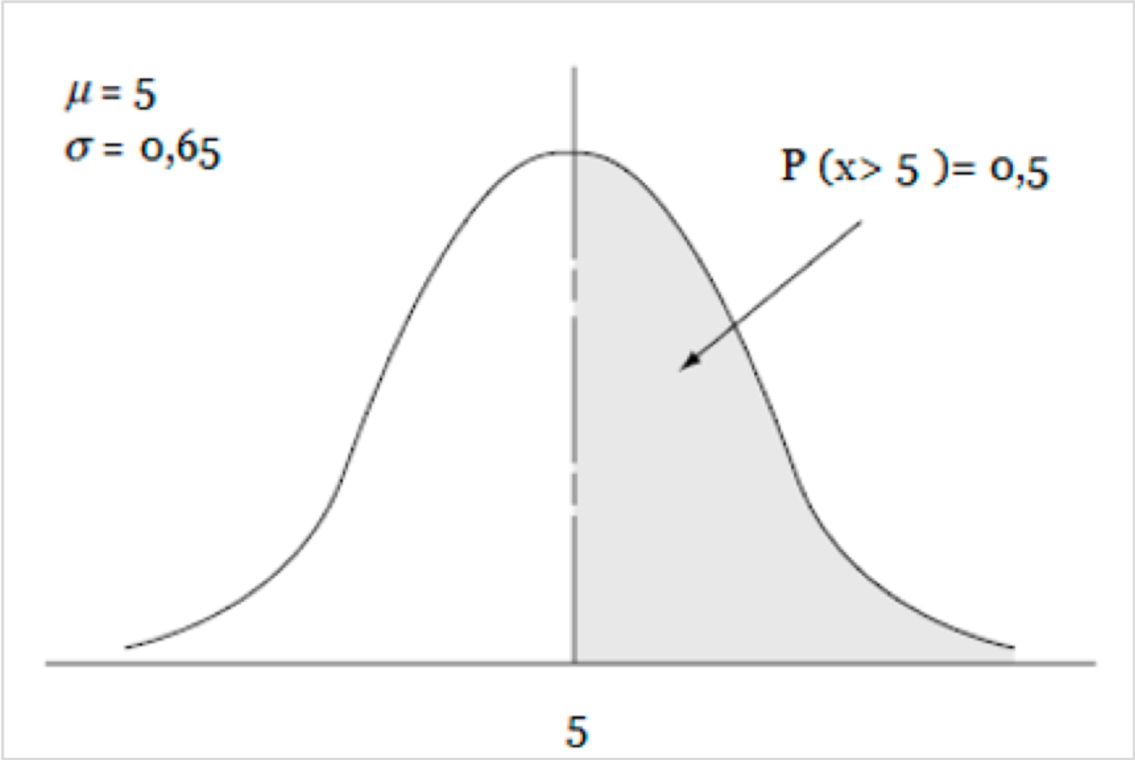
7.4. La tabla de probabilidad normal

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	6.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990

Ejemplo 5. Aplicaciones de la distribución normal

Las horas precisas para la maquetación de unos artículos periodísticos en varios días siguen una distribución normal en la que la media es 5 horas y la desviación 0,65.

A. ¿Cuál es la probabilidad de que un día se hayan necesitados más de 5 horas?



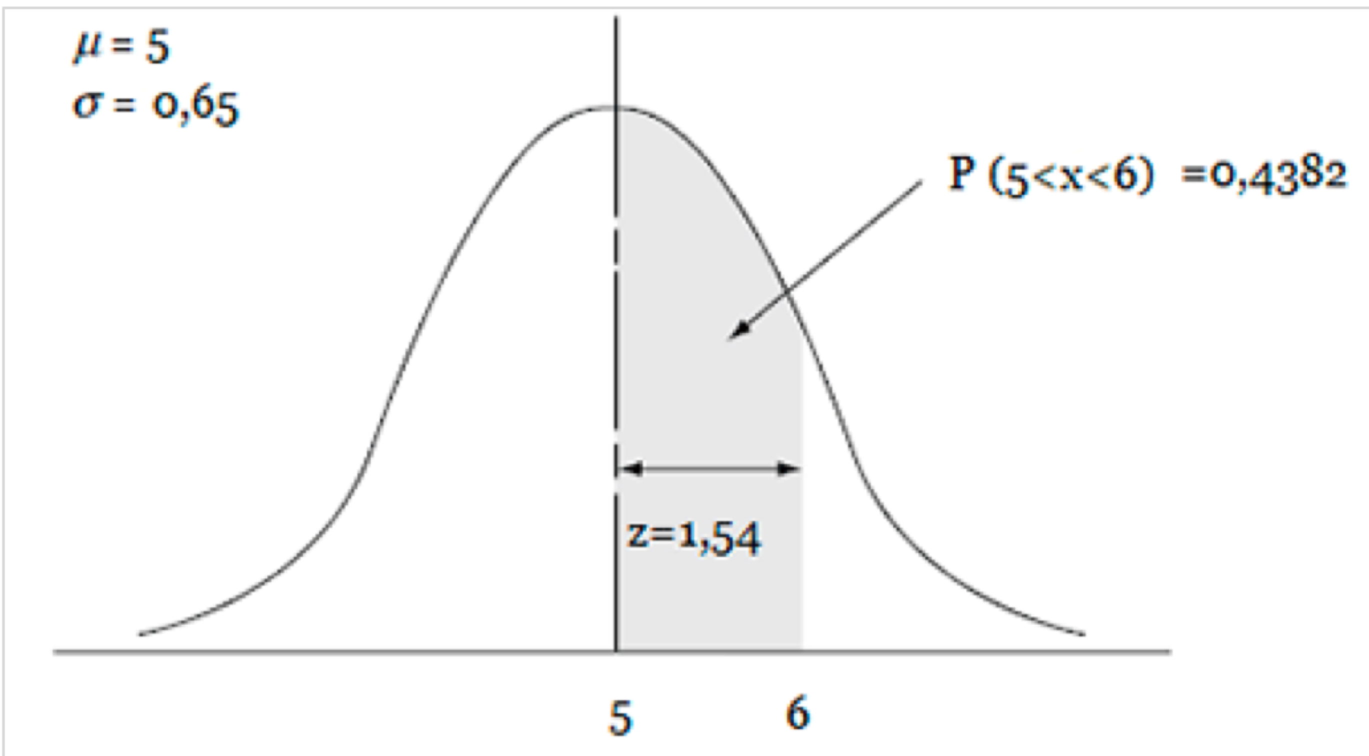
La probabilidad es 0,5 ya que la mistad del área bajo la curva se halla a ambos lados de la media.

B. ¿ Cuál es la probabilidad de que de un día se hayan necesitado entre 5 y 6 horas?

Hay que calcular el valor de Z:

$$z = \frac{x - \mu}{\sigma} = \frac{6 - 5}{0,65} = 1,54\ desviaciones\ estándar$$

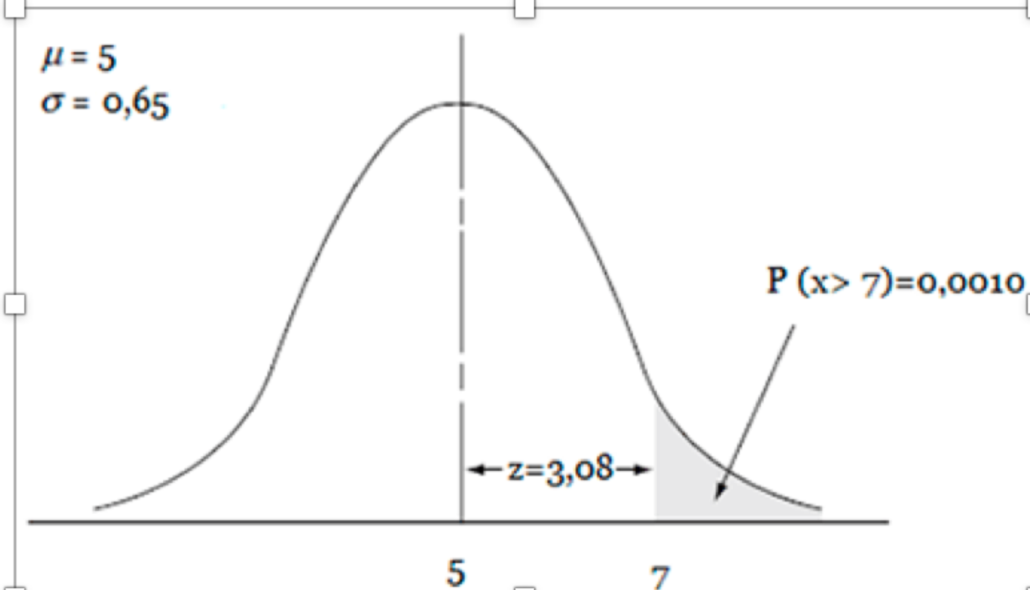
Hay que buscar en la tabla el valor de z=1,54. Para ello miramos en la columna de z hasta la fila con valor 1.5 y hacemos coincidir con la columna 0,04. La probabilidad es 0,4382.



C. ¿Cuál es la probabildad de que un día se hayan requerido 7 horas?

$$z = \frac{x - \mu}{\sigma} = \frac{7 - 5}{0{,}65} = 3{,}08 \; desviaciones \; estándar$$

Comparando el valor de z en la tabla se tiene que la probabilidad es 0,4990. Sin embargo, este dato muestra la probabilidad de que los datos se hallen entre 5 y 7 horas tal y como se ha visto en el apartado B con 6 horas. Tal y como muestra la siguiente figura, la probabilidad de que en un día se hayan requerido más de 7 horas es 0,5 -0,4990, es decir, 0,0010. Lo cual tiene más lógica que pensar que tal probabilidad fuese 0,4990, ya que 7 se aleja bastante de la media.



7.4. La tabla de probabilidad normal

Distribuciones de Muestreo Normales

- **Distribución de probabilidad de las medidas de tendencia central** de un conjunto de muestras, generalmente las medias.
- Cuando llevamos a cabo un muestreo para tomar diferentes muestras de una población, cada una tiene un **valor medio y una desviación estándar**.
- La distribución de todas esas medias es la distribución de muestreo de la media (**error estándar**)
- El error estándar da información de la **magnitud del error y de la precisión** que supone el hecho de tomar ese estadístico a la hora de hacer inferencia en la población

La desviación estándar de una población viene dada por la siguiente expresión:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

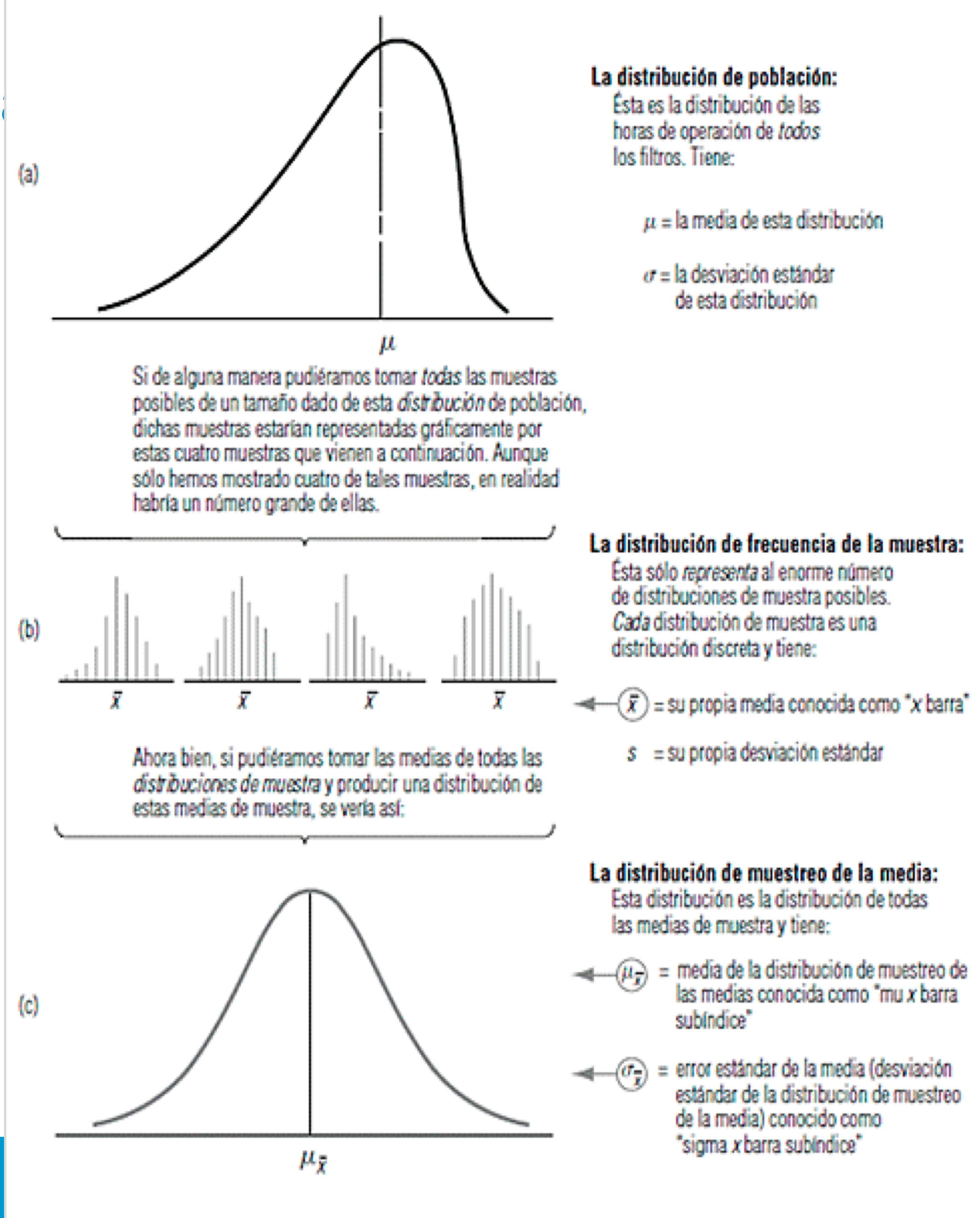
Ejemplo 6. Distinción entre una muestra de una distribución muestral y de una población

Se desea saber las personas que quieren asistir a ver una película el día del estreno. Para ello se eligieron 100 localidades y se entrevistó a 200 personas por localidad, ¿esa muestra es de la población o de una distribución?

Se trata de una muestra de una distribución muestral de las medias de muestras de tamaño 200, obtenida de la población de personas en las localidades.

Si en una población con media μ y desviación estándar σ se toman diferentes muestras aleatorias, cada una tendrá valores diferentes de media y desviación estándar. En la siguiente figura del libro se observan las diferentes distribuciones de unas muestras aleatorias y la distribución de muestreo de la media obtenida a partir de ellas.

7.4. La t

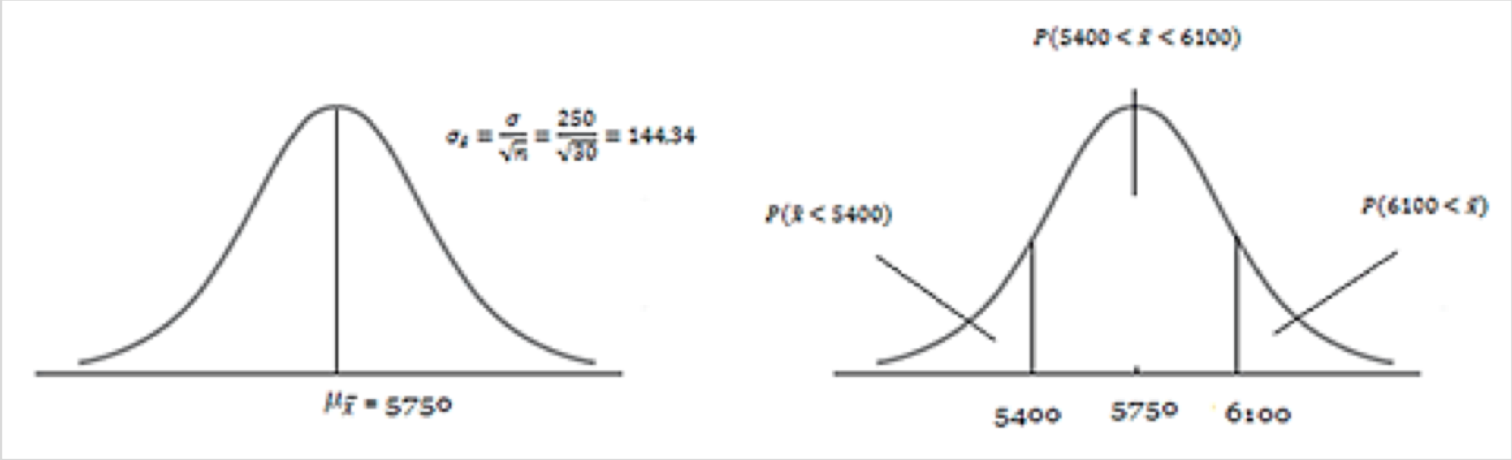

(a)
μ
La distribución de población:
Ésta es la distribución de las horas de operación de *todos* los filtros. Tiene:
μ = la media de esta distribución
σ = la desviación estándar de esta distribución
Si de alguna manera pudiéramos tomar *todas* las muestras posibles de un tamaño dado de esta *distribución* de población, dichas muestras estarían representadas gráficamente por estas cuatro muestras que vienen a continuación. Aunque sólo hemos mostrado cuatro de tales muestras, en realidad habría un número grande de ellas.
(b)
$\bar{x}$
$\bar{x}$
$\bar{x}$
$\bar{x}$
La distribución de frecuencia de la muestra:
Ésta sólo *representa* al enorme número de distribuciones de muestra posibles. *Cada* distribución de muestra es una distribución discreta y tiene:
$\bar{x}$ = su propia media conocida como "*x* barra"
s = su propia desviación estándar
Ahora bien, si pudiéramos tomar las medias de todas las *distribuciones de muestra* y producir una distribución de estas medias de muestra, se vería así:
(c)
$\mu_{\bar{x}}$
La distribución de muestreo de la media:
Esta distribución es la distribución de todas las medias de muestra y tiene:
$\mu_{\bar{x}}$ = media de la distribución de muestreo de las medias conocida como "mu *x* barra subíndice"
$\sigma_{\bar{x}}$ = error estándar de la media (desviación estándar de la distribución de muestreo de la media) conocido como "sigma *x* barra subíndice"

Ejemplo 7. La población de personas que acude a un restaurante de comida rápida diariamente se distribuye de forma normal con una media de 750 y una desviación estándar de 120. Si para hacer un determinado análisis se toman 20 muestras de 15 personas cada una, ¿cuál es la media y la desviación estándar de la distribución muestral?

$$\mu_{\bar{x}} = \mu = 750$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{120}{\sqrt{15}} = 31$$

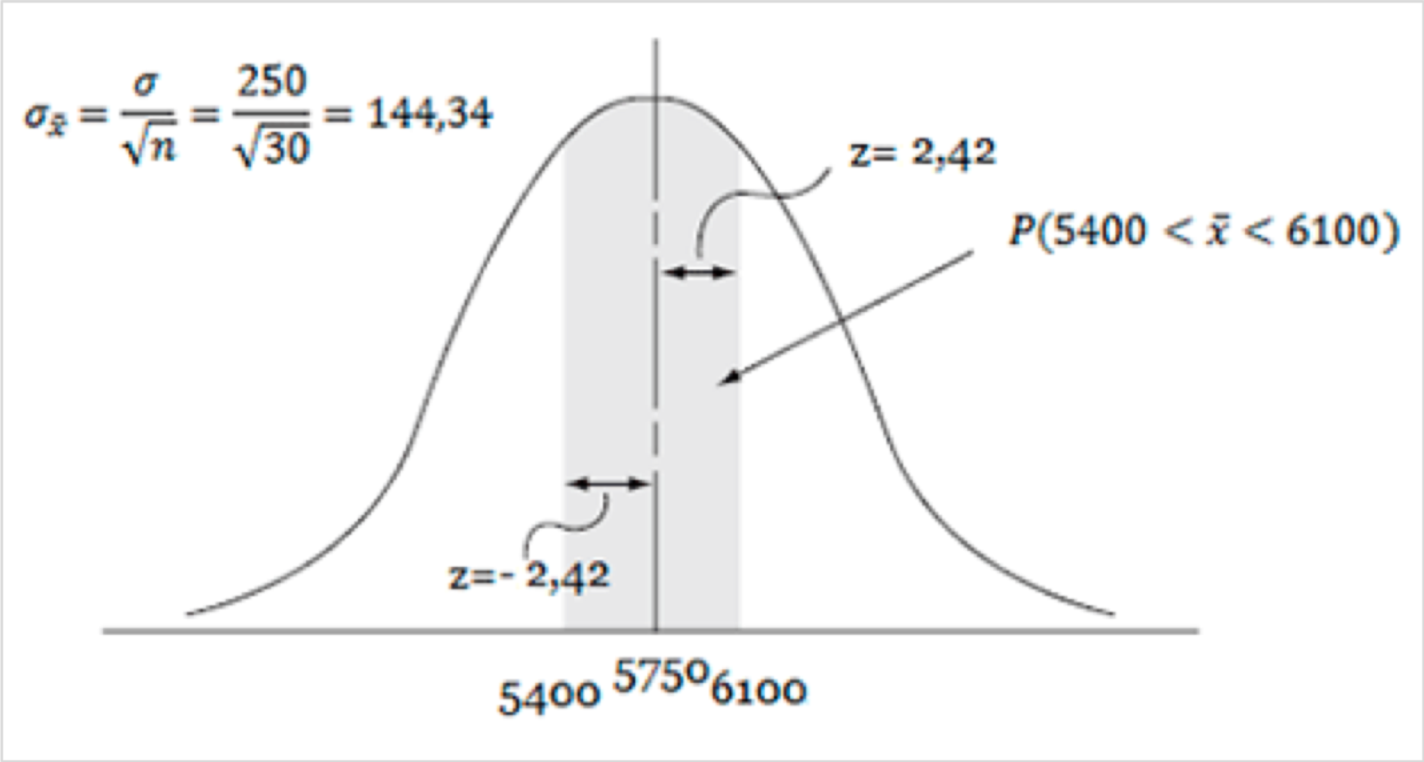
Ejemplo 8. Un periodista quiere escribir un artículo sobre la cantidad de residuos reciclados en una comunidad en la que la media es 5750 kg y la desviación estándar 250. Para estudiar cómo se distribuye el reciclaje en las diferentes regiones toma muestras de 30 empresas de reciclaje y considera que las medias muestrales serán buenas estimaciones de la media poblacional si difieren de ella en ± 350 toneladas. ¿Cuál es la probabilidad de que la media muestral se halle a ese valor de la media poblacional?



Lo que se pretende saber es la probabilidad de que los valores de la media estén entre 5400 y 6100.

$$z = \frac{x-\mu}{\sigma} = \frac{5400-5750}{144,34} = -2,42 \qquad z = \frac{x-\mu}{\sigma} = \frac{6100-5750}{144,34} = 2,42$$

Comparando en la tabla de distribucion de probabilidad normal estándar resulta que la probabilidad total es 0,4922+0,4922=0,9844. Tal porcentaje es bastate elevado por lo que la estimación es adecuada.



7.5. La regla 68/95/99,7

- El 68% aproximadamente de los datos caen a ±1 desviacion esátandar de la media.
- El 95% aproximadamente de los datos caen a ±2 desviaciones estándar de la media.
- El 99,7 % aproximadamente de los datos caen a ±3 desviaciones estándar de la media.

Ejemplo 9. Las alturas de unos jugadores de baloncesto siguen una distribución normal con media 1,77 y desviación estándar 0,07.

A. ¿Qué porcentaje de jugadores mide más 1,91 m?

- El intervalo para una desviacion estándar es:
  1,77 -0,07 hasta 1,77+0,07, es decir, 1,70-1,84
- El intervalo para dos desviaciones estándar es:
  1,77-0,14 hasta 1,77+0,141, es decir, 1,63-1,91
- El intervalo para tres desviaciones estándar es:
  1,77-0,21 hasta 1,77+0,21, es decir, 1,56-1,98

El valor 1,91 está en el límite superior de dos desviaciones estándar. Este intervalo recoge al 95% de los datos, por lo que los datos que están fuera suponenen el restante 5%. La mitad de este 5% es para los datos menores a los del intervalo y la mitad para los mayores, por lo que el porcentaje de que la altura sea superior a 1,91 m es 2,5%.

UNIVERSIDAD INTERNACIONAL DE LA RIOJA

www.unir.net