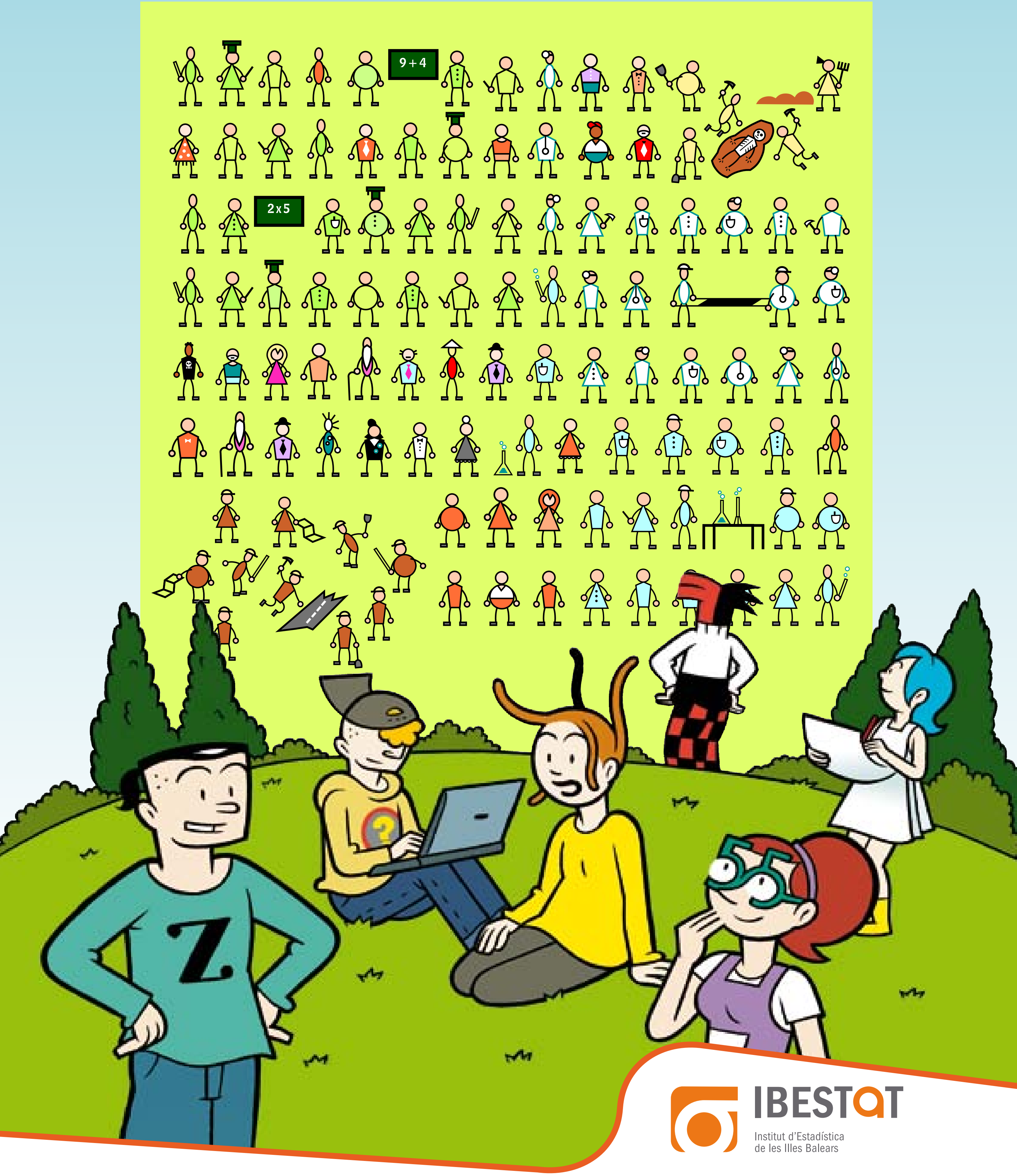
DADOS Y DATOS III

El paso de la incertidumbre al riesgo



IBESTAT
Institut d'Estadística de les Illes Balears

DADOS Y DATOS III

El paso de la incertidumbre al riesgo

Govern de les Illes Balears

IBESTAT
Institut d'Estadística
de les Illes Balears

© Edición: **Institut d'Estadística de les Illes Balears (IBESTAT)**
Dirección del proyecto: **Andreu Sansó Rosselló**
Coordinación general: **Damià Perelló Femenia y Sara Fernández Vázquez**
C/ de Sant Sebastià, 1
07001 Palma (Mallorca)
Teléfono: 971 784 575
Fax: 971 784 579

Autor: **Javier Cubero**

Gestión y producción: **inrevés SL**
Ilustraciones y maquetación: **Alex Fito**
Coordinación y guión adaptado: **Pere Joan**

Colección: **Estadística al carrer. Volumen 3**
Título: **Dados y datos III. El paso de la incertidumbre al riesgo**
Núm. **IBESTAT: 4/2008**
Depósito legal:

Impresión: **Jorvich**
Fecha de edición: **2010**

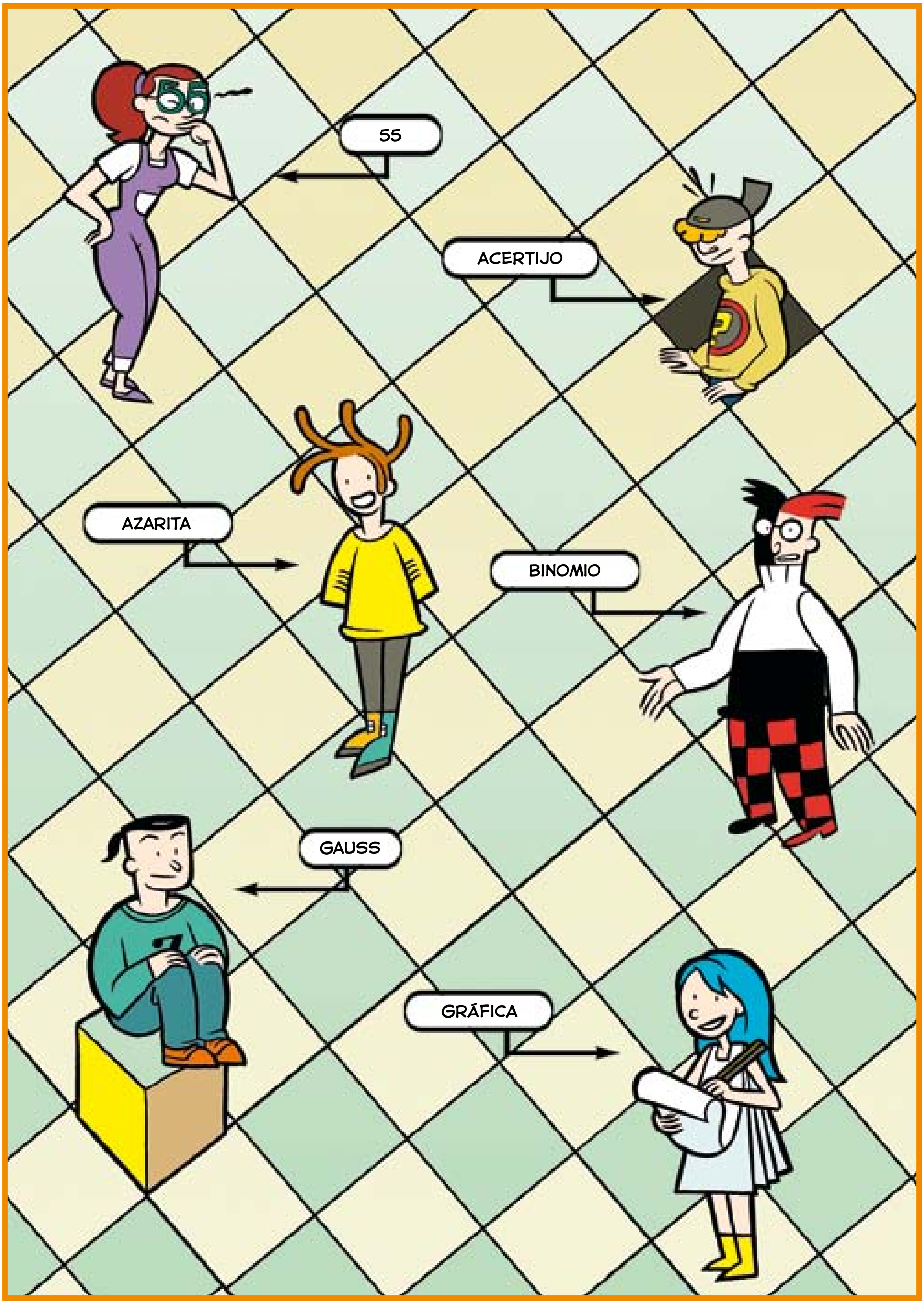
Con la constitución del Institut d'Estadística de les Illes Balears (IBESTAT), este 2008, como entidad autónoma, se ha querido dar un paso muy importante en todo lo que debe ser la vertebración de un verdadero sistema estadístico para nuestra comunidad autónoma. También ha sido un momento de análisis y de reflexión sobre la labor que ya se había llevado a cabo en etapas anteriores y que, por su calidad y valía, era necesario recuperar y proyectar hacia el futuro. Este ha sido el caso de la colección Dados y datos, que con Javier Cubero como autor, ha sido capaz de encontrar una formulación muy pedagógica para hacer llegar los grandes conceptos estadísticos a amplios sectores de la población, entre los cuales cabe mencionar a los más jóvenes.

El formato amable y atractivo que supone el uso del cómic como soporte lo hace prácticamente único en su especialidad, pero no por ello falto de rigor ni del nivel que requiere una publicación de carácter científico como ésta. Esto ha llevado a que las dos ediciones anteriores, que se materializaron bajo los auspicios de los directores generales de Economia Antoni Monserrat (volumen I) y Maria Marquès (volumen II), hayan tenido que ser reeditadas y completadas con un tercer volumen que cerrará una trilogía muy completa por lo que respecta al conocimiento de los principios estadísticos. Todo este proceso se enmarca en la voluntad de fomentar la estadística como disciplina útil para el conocimiento de la realidad que nos rodea a partir de elementos que aparentemente son sencillos pero cuyo alcance formativo es grande.

Para terminar, quisiera agradecer la colaboración de todas las personas que han participado en la edición de este material didáctico, tanto por lo que respecta a los creativos y dibujantes como a los técnicos. Igualmente, me gustaría animar a todo el mundo para que se acerque a estas publicaciones y así descubrir un mundo lleno de conocimientos que sin duda nos ayudarán a comprender mejor nuestra realidad de una forma más apasionante y racional a la vez.

Andreu Sansó Rosselló

Director del IBESTAT


55
ACERTIJO
AZARITA
BINOMIO
GAUSS
GRÁFICA

Capítulo 1

GEORGES LOUIS LECLERC

Conde de Buffón (1773)

Montbard, Borgoña,1707 – París, 1788

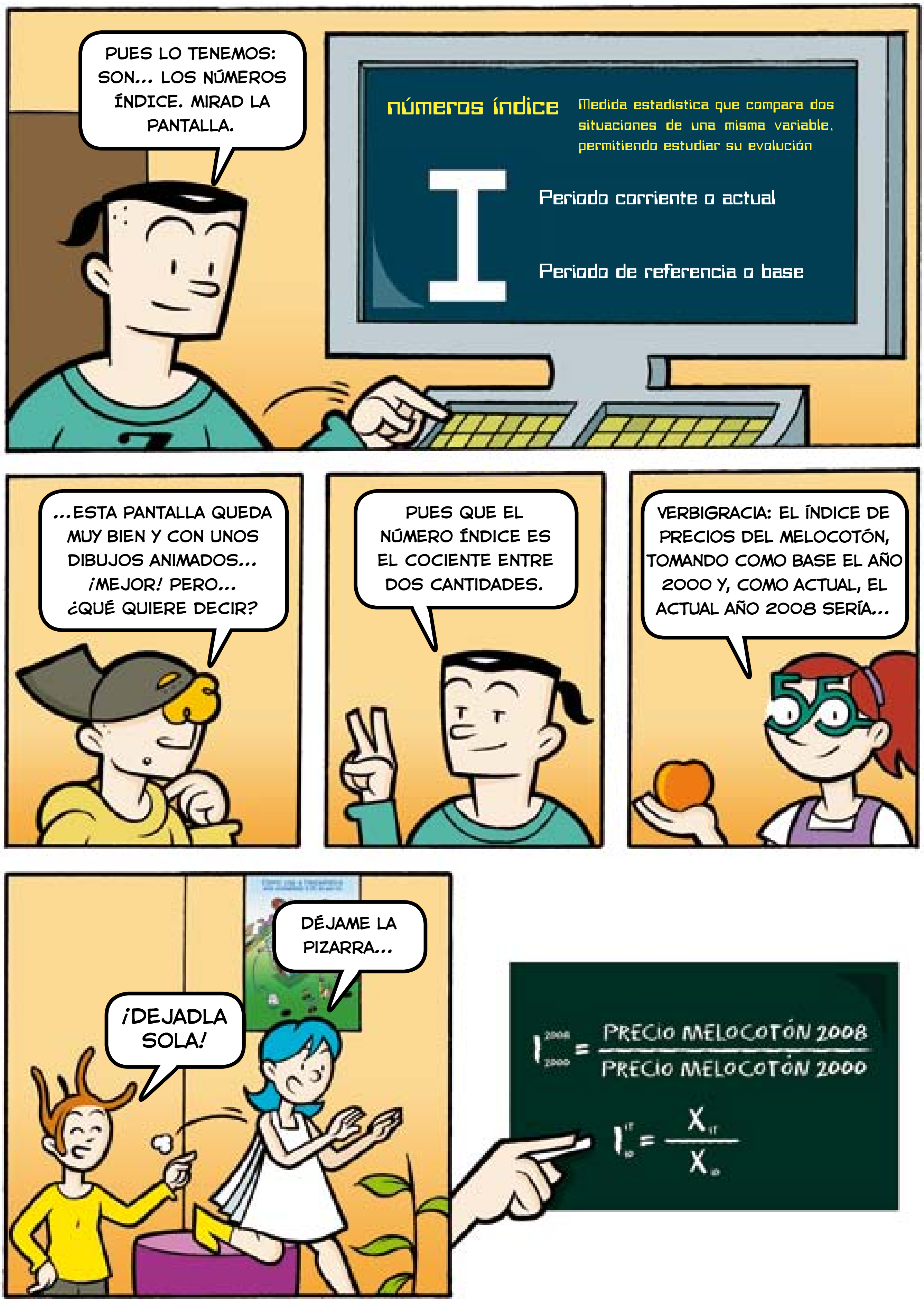
Matemático y naturalista francés.

Estudioso de la mecánica, la astronomía, medicina..., la teoría de los números, el cálculo, la geometría y **la probabilidad**.

Se hizo famoso en esta última materia por su curioso experimento de determinar muchas cifras decimales en el valor del número π tratándolo como una experiencia de **probabilidad geométrica** conocida como la aguja de Bufón.


YA COMENZAMOS, ¡AHORA UN PROBLEMA! DEBEMOS REALIZAR UN TRABAJO SOBRE: NÚMEROS ÍNDICES.
ESO... ¿NO ES LO DEL I.P.C. Y TODOS ESOS NUMERACOS ECONÓMICOS QUE HOY DÍA VAN TAN MAL?
PUES VAYA TIEMPO PARA MANDARNOS EL TRABAJO, EN OTRAS FECHAS HUBIERA SIDO UN EXITAZO.
O SEA, SI MI ABUELO COMPRABA UNA ENSAIMADA A 6 PESETAS...
¡¿QUÉ?!
UN POCO MENOS DE 0'04€. LO REPITO: MENOS DE CUATRO CÉNTIMOS.
...Y LO COMPARAMOS CON LO QUE NOS CUESTA HOY...

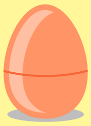

BUENO, ALGO DE ESO Y DE MUCHAS COSAS MÁS QUE COMPARAMOS; PUES FIJAOS, PODRÍAMOS PENSAR LO SIGUIENTE:
¿CUÁNTOS DÍAS U HORAS O MINUTOS TENÍA QUE TRABAJAR MI ABUELO PARA COMPRAR UNA ENSAIMADA? ...¿Y CUÁNTOS TENDRÍA QUE TRABAJAR YO, HOY DÍA? Y ¿SERÍA UNA ENSAIMADA DE IGUALES CARACTERÍSTICAS?
SIN QUERER HEMOS ROZADO LOS CONCEPTOS DE DEFLACTACIÓN, PRECIO EN UNIDADES CONSTANTES, ÍNDICE DE PRECIO, HASTA MUY DE LEJOS EL DE PRECIO HEDÓNICO.
HAS PASADO MALA NOCHE... ¿VERDAD?
BUENO, LO QUE SÍ VEO ES QUE TENEMOS QUE INVENTAR ALGO GENERAL QUE DEFINA ESAS COMPARACIONES Y MÁS QUE PUDIERAN SURGIR.
Y TENDRÍA QUE TENER TRATAMIENTO MATEMÁTICO, PARA SU APLICACIÓN CORRECTA.



PUES LO TENEMOS: SON... LOS NÚMEROS ÍNDICE. MIRAD LA PANTALLA.
números índice
Medida estadística que compara dos situaciones de una misma variable, permitiendo estudiar su evolución
I
Periodo corriente o actual
Periodo de referencia o base
...ESTA PANTALLA QUEDA MUY BIEN Y CON UNOS DIBUJOS ANIMADOS... ¡MEJOR! PERO... ¿QUÉ QUIERE DECIR?
PUES QUE EL NÚMERO ÍNDICE ES EL COCIENTE ENTRE DOS CANTIDADES.
VERBIGRACIA: EL ÍNDICE DE PRECIOS DEL MELOCOTÓN, TOMANDO COMO BASE EL AÑO 2000 Y, COMO ACTUAL, EL ACTUAL AÑO 2008 SERÍA...
DÉJAME LA PIZARRA...
¡DEJADLA SOLA!
$I_{2000}^{2008} = \frac{\text{PRECIO MELOCOTÓN 2008}}{\text{PRECIO MELOCOTÓN 2000}}$
$I_0^t = \frac{X_t}{X_0}$

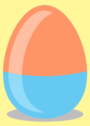
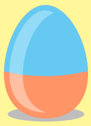

OS ACLARO QUE AUNQUE PARA OPERAR LO HAREMOS COMO INDICA LA PIZARRA, GENERALMENTE PARA SU EXPOSICIÓN SE MULTIPLICA EL RESULTADO POR CIEN.
ENTONCES LO DAS EN TANTOS POR CIENTO.
¡OLÉ! ¡EUREKA!
¡"UNO"!
HE ESTADO CON GRÁFICA EN LA BIBLIOTECA Y HEMOS OBSERVADO ALGUNAS PROPIEDADES DE LOS NÚMEROS ÍNDICE.
HEMOS ELABORADO UNOS GRÁFICOS QUE PUEDEN SERVIRNOS PARA EL TRABAJO. ¡MIRAD!
Daus i dades

IDENTIDAD

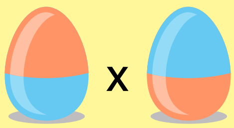
 $= 1$

$$I_t^t = \frac{x_{it}}{x_{it}} = 1$$

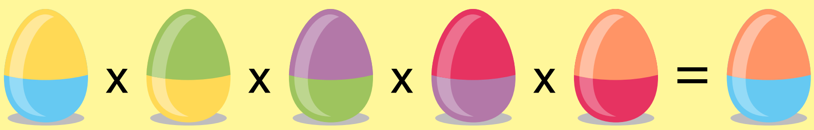
INVERSIÓN

 $= \frac{1}{}$ 

$$I_0^t = \frac{1}{I_t^0} \qquad I_2^7 = \frac{1}{I_7^2}$$

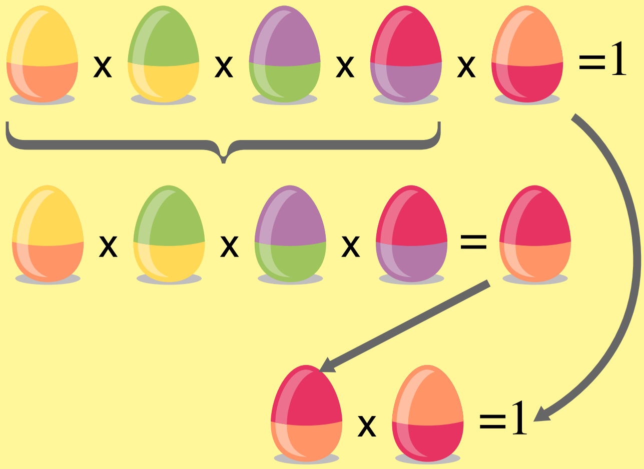
 $= 1$

CÍCLICA



$$I_0^1 \times I_1^2 \times I_2^3 \times I_3^4 = I_0^4 \qquad I_0^1 \times I_1^3 \times I_0^{10} \times I_{10}^{13} = I_0^{13} \qquad I_0^1 \times I_1^2 \times \frac{1}{I_5^2} \times I_5^8 = I_0^8$$

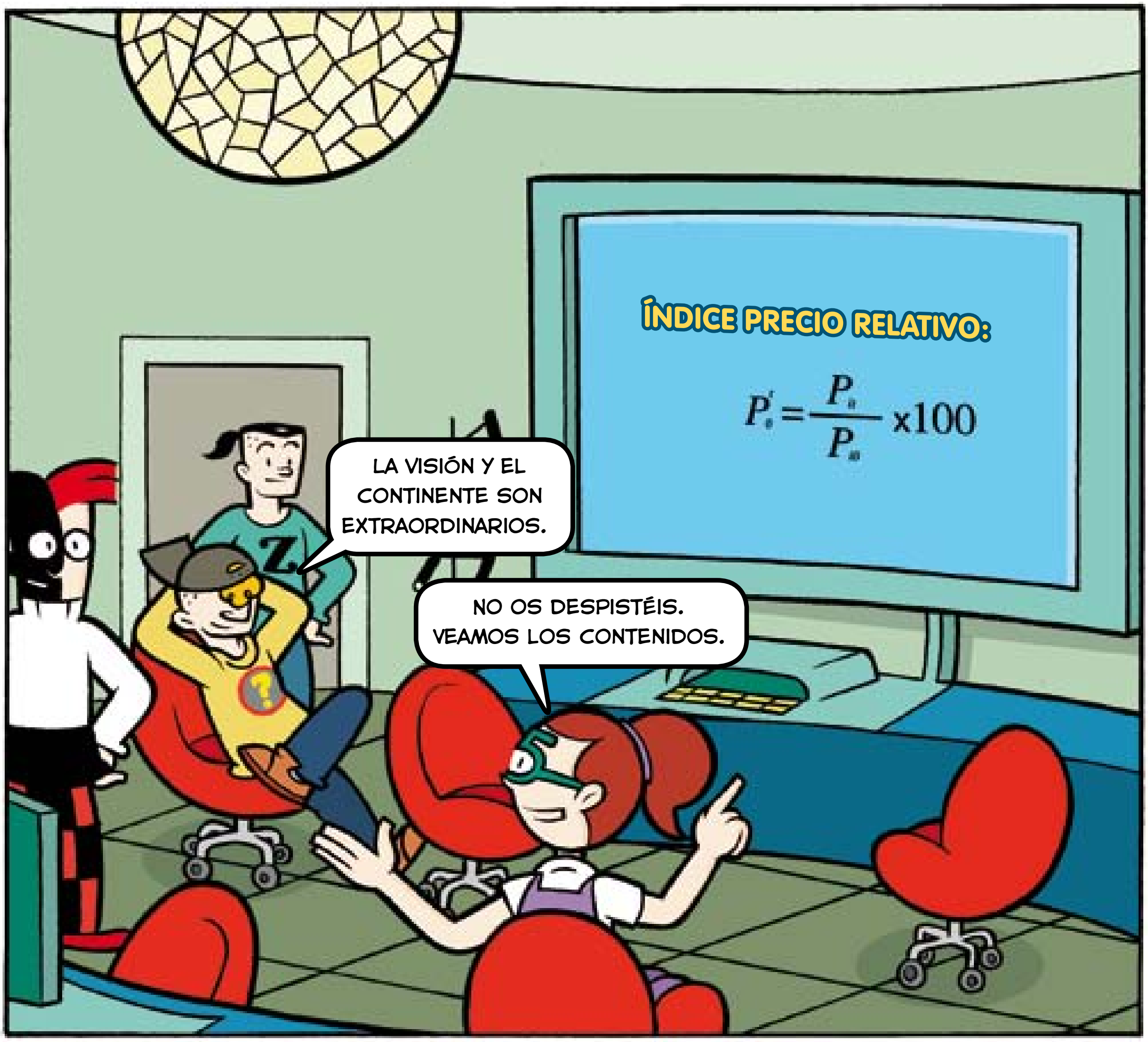
CIRCULAR



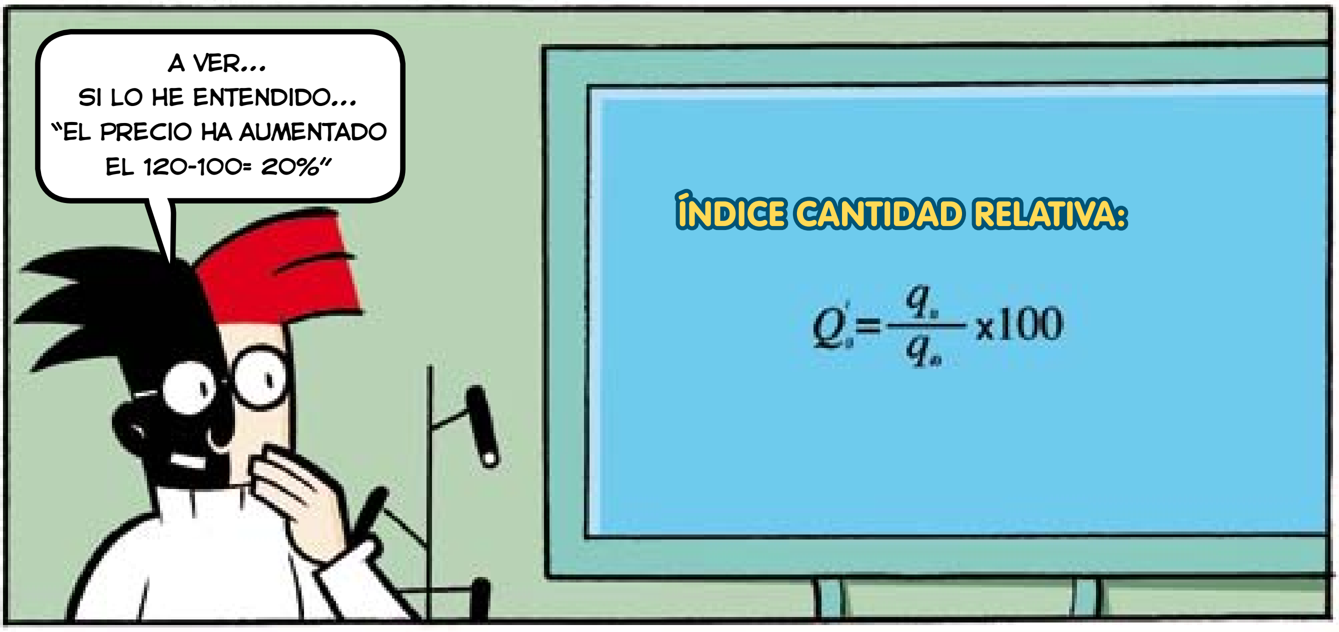
$$I_0^1 \times I_1^2 \times I_2^3 \times I_3^0 = 1$$

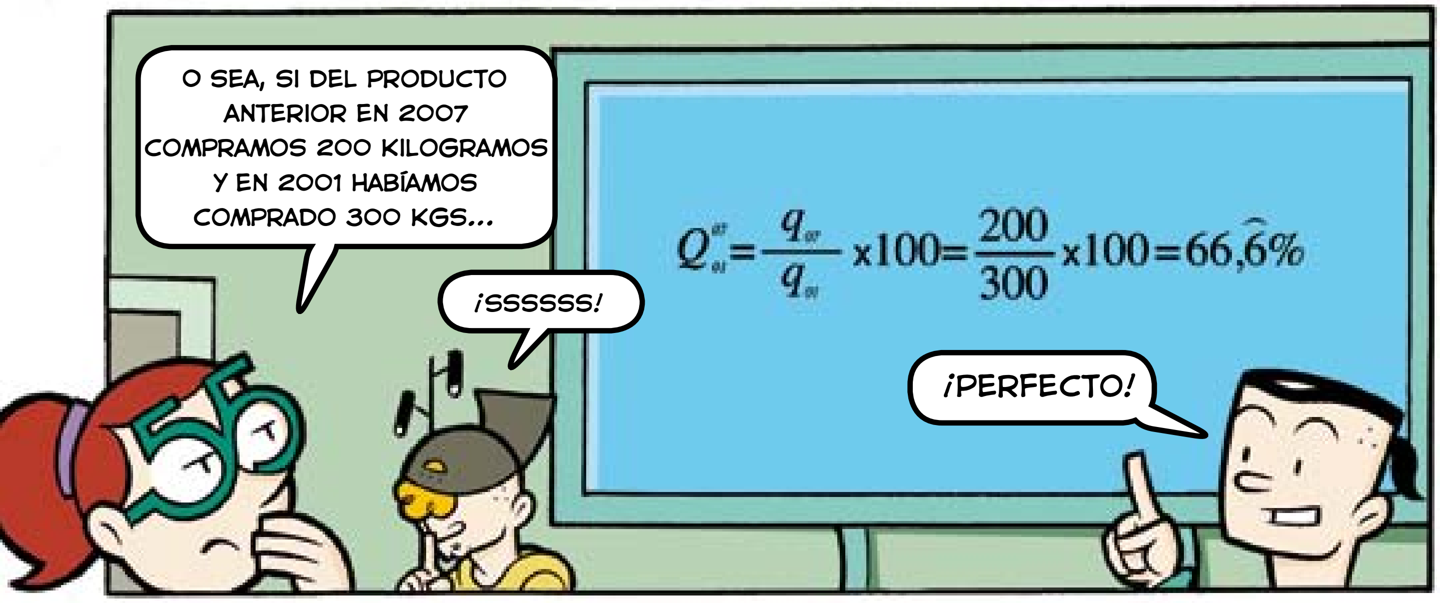
$$I_1^3 \times \frac{1}{I_5^3} \times I_5^1 = 1$$

UNA SEMANA DESPUÉS



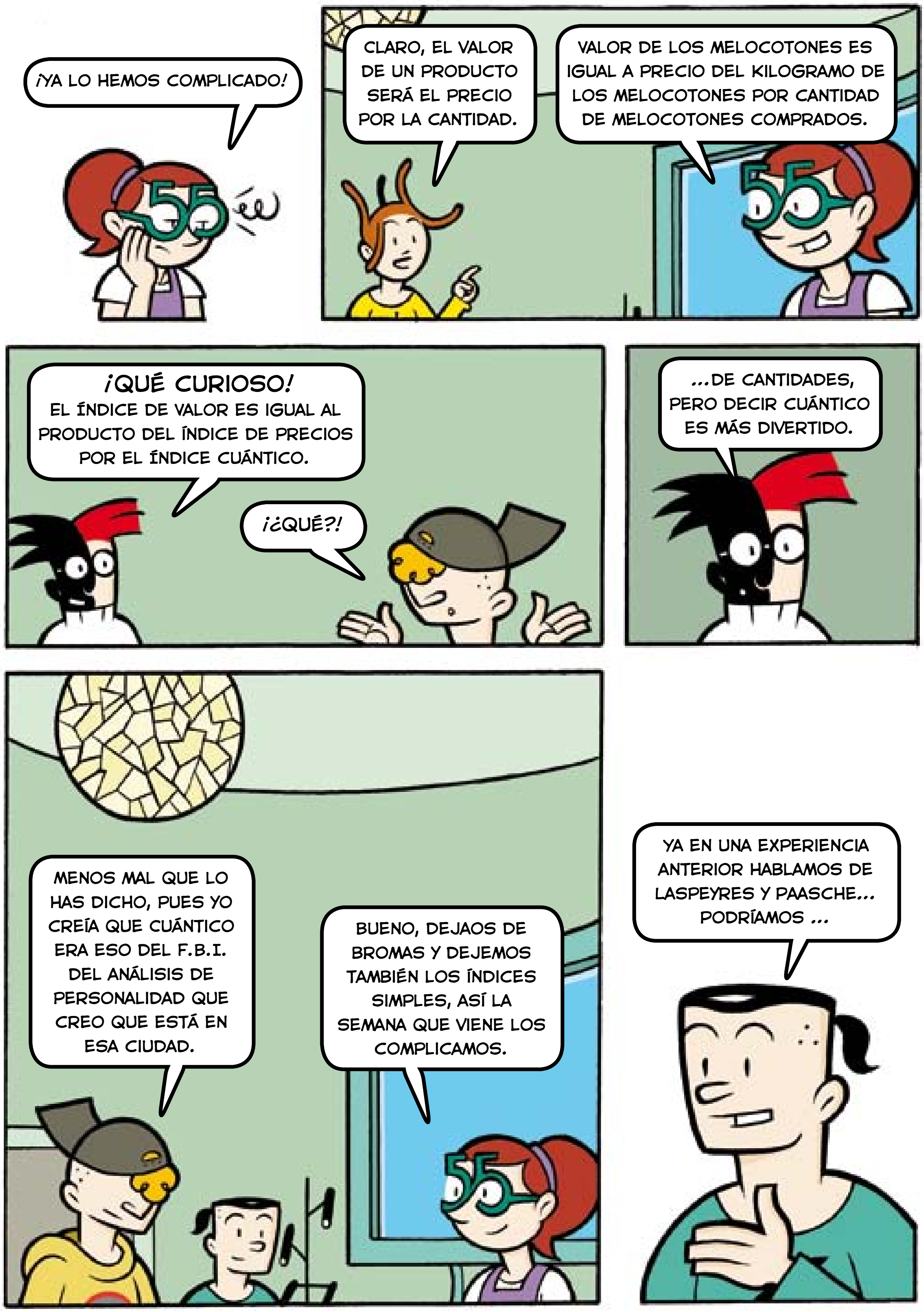
$$P_{01}^{07} = \frac{P_{07}}{P_{01}} x100 = \frac{30}{25} x100 = 120\%$$



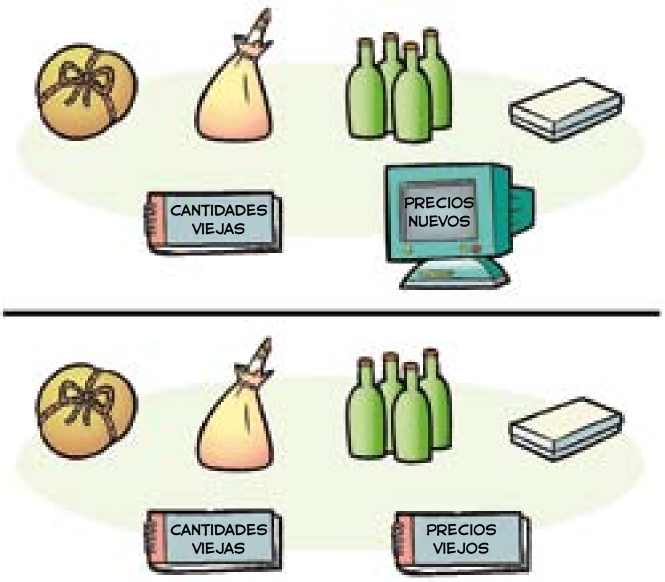


ÍNDICE VALOR RELATIVO:

$$V'_0 = \frac{V_{it}}{V_{i0}} \times 100 = \frac{p_{it}\, q_{it}}{p_{i0}\, q_{i0}} \times 100 = \left(\frac{p_{it}}{p_{i0}}\right) \times \left(\frac{q_{it}}{q_{i0}}\right) \times 100 = P'_0 \times Q'_0 \times 100$$

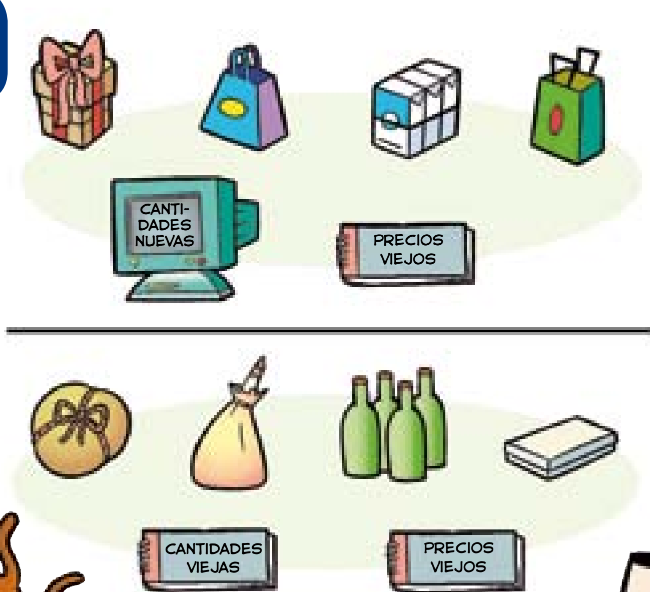

¡YA LO HEMOS COMPLICADO!
CLARO, EL VALOR DE UN PRODUCTO SERÁ EL PRECIO POR LA CANTIDAD.
VALOR DE LOS MELOCOTONES ES IGUAL A PRECIO DEL KILOGRAMO DE LOS MELOCOTONES POR CANTIDAD DE MELOCOTONES COMPRADOS.
¡QUÉ CURIOSO! EL ÍNDICE DE VALOR ES IGUAL AL PRODUCTO DEL ÍNDICE DE PRECIOS POR EL ÍNDICE CUÁNTICO.
¡¿QUÉ?!
...DE CANTIDADES, PERO DECIR CUÁNTICO ES MÁS DIVERTIDO.
MENOS MAL QUE LO HAS DICHO, PUES YO CREÍA QUE CUÁNTICO ERA ESO DEL F.B.I. DEL ANÁLISIS DE PERSONALIDAD QUE CREO QUE ESTÁ EN ESA CIUDAD.
BUENO, DEJAOS DE BROMAS Y DEJEMOS TAMBIÉN LOS ÍNDICES SIMPLES, ASÍ LA SEMANA QUE VIENE LOS COMPLICAMOS.
YA EN UNA EXPERIENCIA ANTERIOR HABLAMOS DE LASPEYRES Y PAASCHE... PODRÍAMOS ...

ÍNDICE LASPEYRES DE PRECIOS
L_p



$$L_p = \frac{\textit{cantidades viejas} \times \textit{precios nuevos}}{\textit{cantidades viejas} \times \textit{precios viejos}}$$

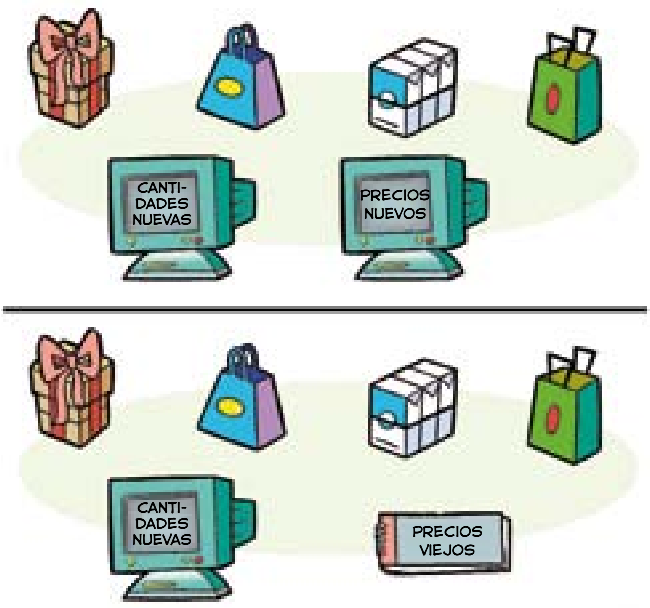
ÍNDICE LASPEYRES CUÁNTICO
L_Q



$$L_q = \frac{\textit{precios viejos} \times \textit{cantidades nuevas}}{\textit{precios viejos} \times \textit{cantidades viejas}}$$

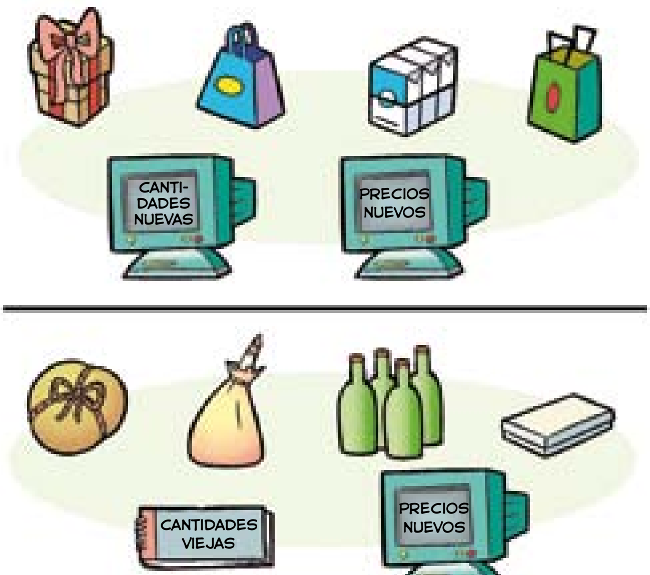
ÍNDICE PAASCHE DE PRECIOS
P_p

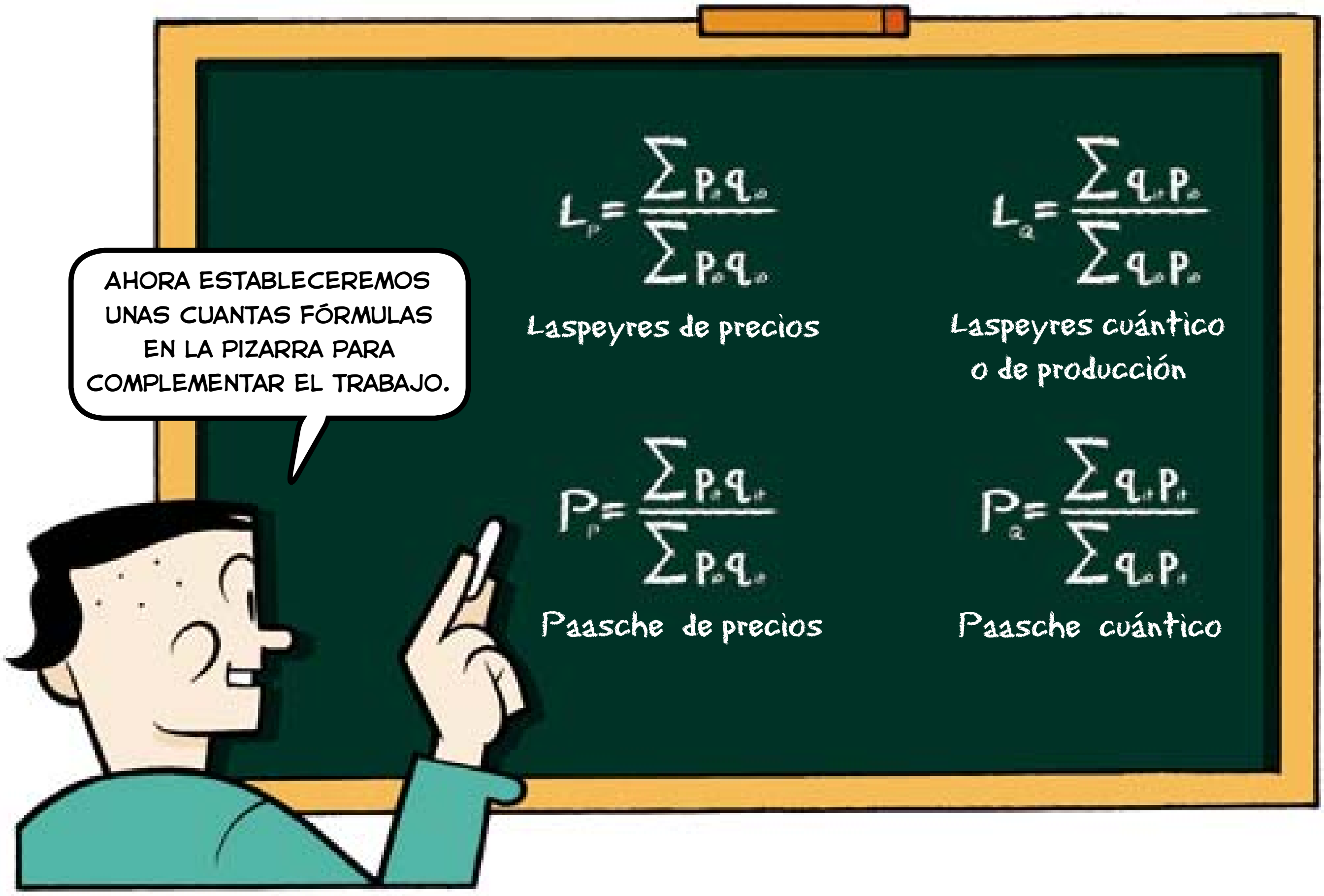
$$P_p = \frac{\textit{cantidades nuevas} \times \textit{precios nuevos}}{\textit{cantidades nuevas} \times \textit{precios viejos}}$$

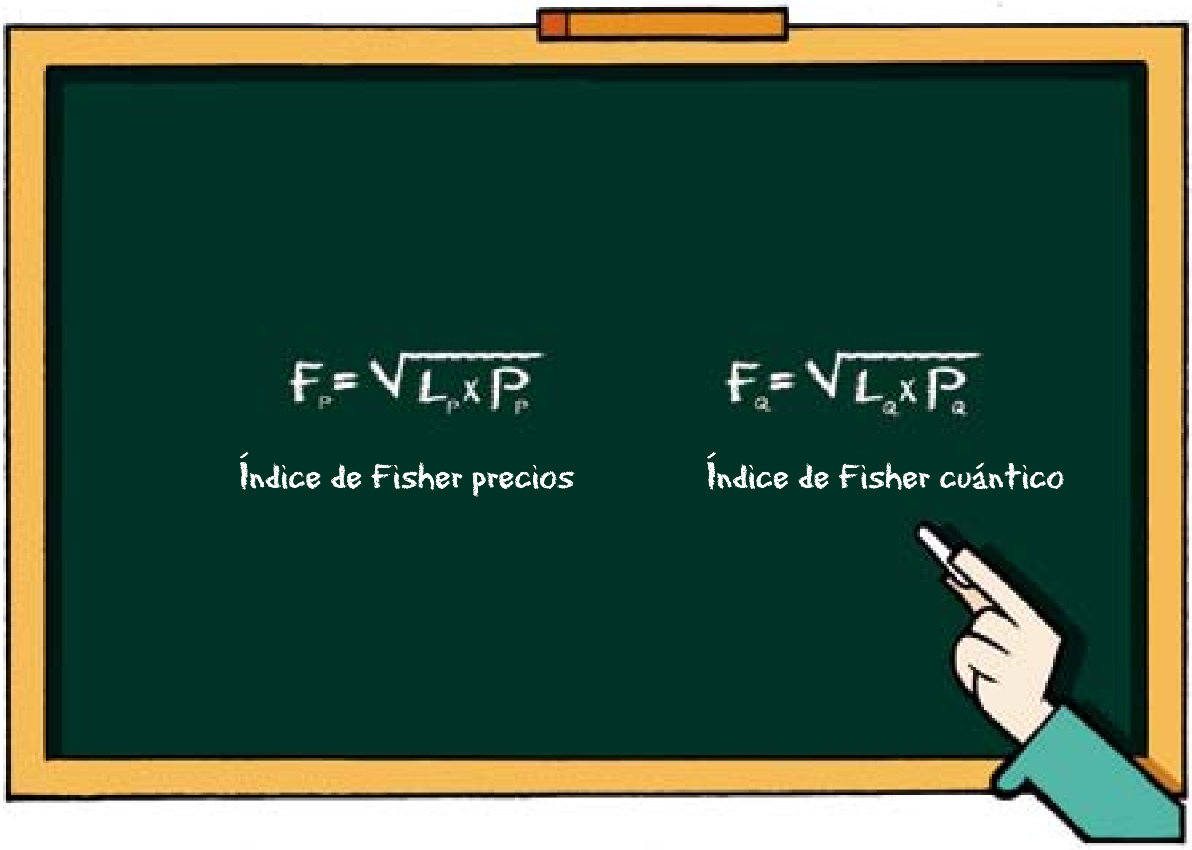


ÍNDICE PAASCHE CUÁNTICO
P_Q

$$P_q = \frac{\textit{precios nuevos} \times \textit{cantidades nuevas}}{\textit{precios nuevos} \times \textit{cantidades viejas}}$$



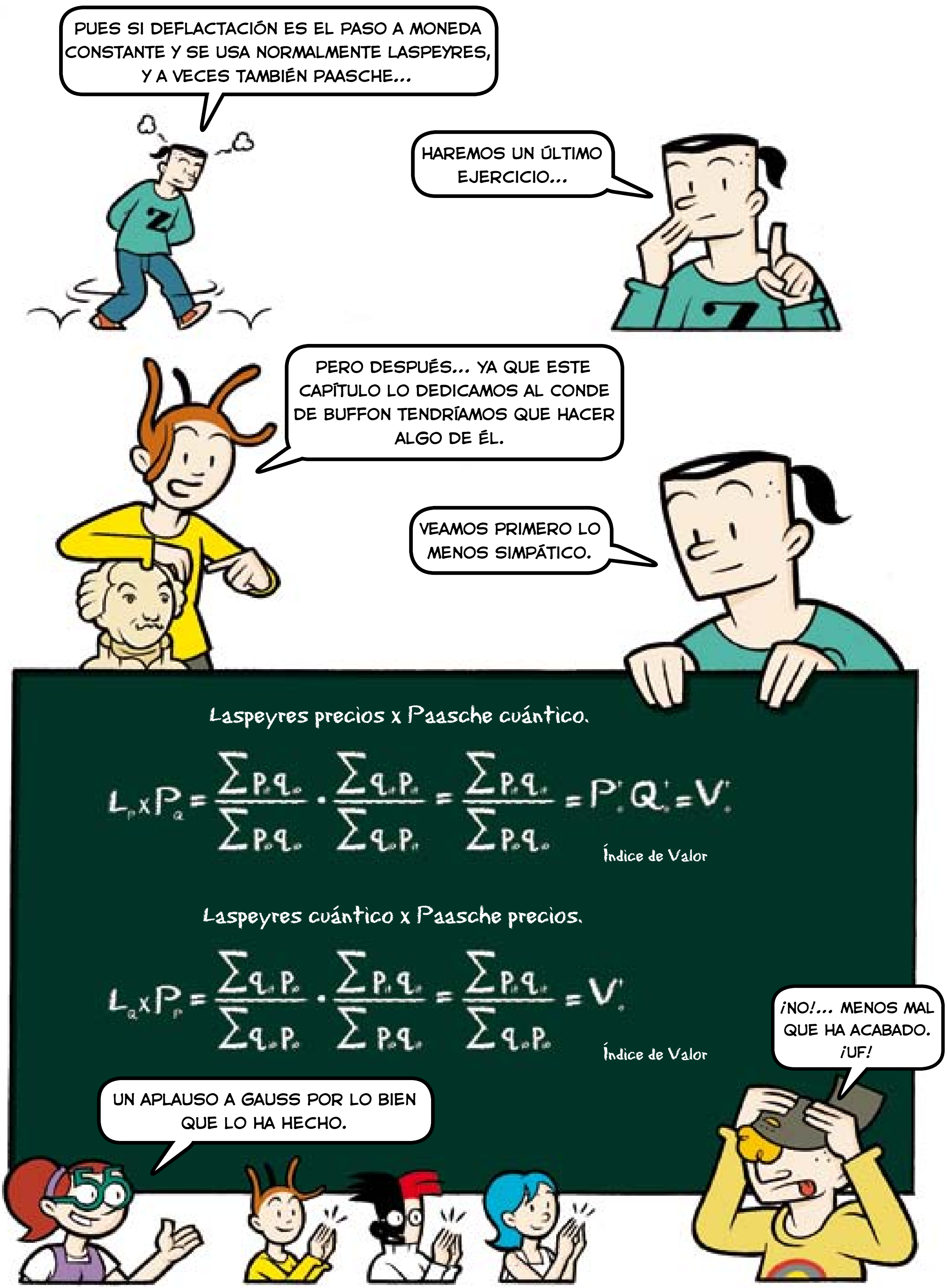

AHORA ESTABLECEREMOS UNAS CUANTAS FÓRMULAS EN LA PIZARRA PARA COMPLEMENTAR EL TRABAJO.
$L_p = \frac{\sum p_1 q_0}{\sum p_0 q_0}$
Laspeyres de precios
$L_q = \frac{\sum q_1 p_0}{\sum q_0 p_0}$
Laspeyres cuántico o de producción
$P_p = \frac{\sum p_1 q_1}{\sum p_0 q_1}$
Paasche de precios
$P_q = \frac{\sum q_1 p_1}{\sum q_0 p_1}$
Paasche cuántico


$F_p = \sqrt{L_p \times P_p}$
Índice de Fisher precios
$F_q = \sqrt{L_q \times P_q}$
Índice de Fisher cuántico


GAUSS, TÚ DIJISTE AL PRINCIPIO ALGO DE "EXPLOSIÓN" O "DEFLAGRACIÓN".
¡CÓMO TE GUSTA HACER EL BURRO!

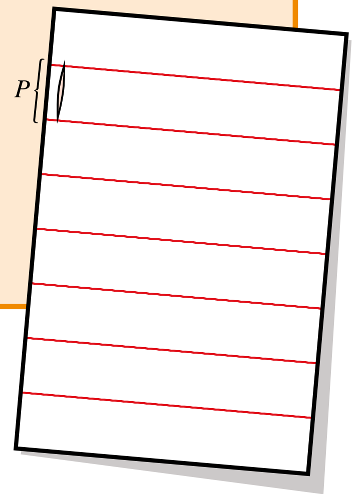

JA, JA, JA
HABLASTE DE DEFLACTACIÓN, Y CREO QUE INTERESA SABER PASAR DE MONEDA CORRIENTE A MONEDA CONSTANTE.


O SEA:
MONEDA CORRIENTE, LA DE CADA PERIODO.
MONEDA CONSTANTE, LA QUE RIGE EN EL PERIODO ESTUDIADO.


PUES SI DEFLACTACIÓN ES EL PASO A MONEDA CONSTANTE Y SE USA NORMALMENTE LASPEYRES, Y A VECES TAMBIÉN PAASCHE...
HAREMOS UN ÚLTIMO EJERCICIO...
PERO DESPUÉS... YA QUE ESTE CAPÍTULO LO DEDICAMOS AL CONDE DE BUFFON TENDRÍAMOS QUE HACER ALGO DE ÉL.
VEAMOS PRIMERO LO MENOS SIMPÁTICO.
Laspeyres precios x Paasche cuántico.
$L_p \times P_q = \frac{\sum p_1 q_0}{\sum p_0 q_0} \cdot \frac{\sum q_1 p_1}{\sum q_0 p_1} = \frac{\sum p_1 q_1}{\sum p_0 q_0} = P_0^1 Q_0^1 = V_0^1$
Índice de Valor
Laspeyres cuántico x Paasche precios.
$L_q \times P_p = \frac{\sum q_1 p_0}{\sum q_0 p_0} \cdot \frac{\sum p_1 q_1}{\sum p_0 q_1} = \frac{\sum p_1 q_1}{\sum q_0 p_0} = V_0^1$
Índice de Valor
¡NO!... MENOS MAL QUE HA ACABADO. ¡UF!
UN APLAUSO A GAUSS POR LO BIEN QUE LO HA HECHO.

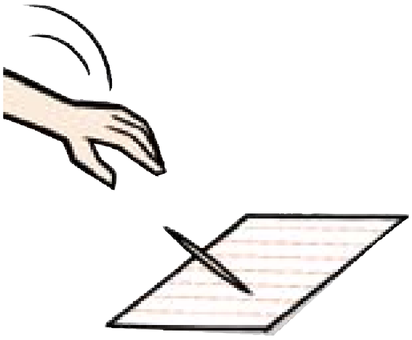
LA AGUJA DE BUFFON

El número π es frecuente e importante en realizaciones estadísticas y el Conde de Buffon, mediante un experimento que de forma simple y sencilla exponemos a continuación, fue uno de los muchos que dedicaron esfuerzos a la tarea de descubrir cuantas más cifras, mejor, de un número que tiene infinitas.



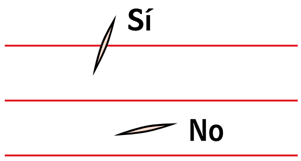
Experimento:

Tomemos una hoja de papel, tracemos en ella ocho particiones iguales, separadas por segmentos como indica la figura, y obtengamos un palillo con longitud "**p**" **igual a la distancia entre los segmentos**, ya sea recortándolo o midiendo primero el palillo y después las divisiones de la hoja.



Realización:

Lancemos el palillo sobre la hoja, cuantas más veces mejor (lanzamientos suaves) y anotemos el número total de lanzamientos al que llamaremos **T**. Vayamos al mismo tiempo contando las veces que de estos lanzamientos la aguja corta a las rayas rojas; a este conteo le llamaremos **C**.



Número de lanzamientos T

Número de cortes C

El conteo puedes hacerlo como lo indicamos en una de nuestras primeras experiencias.

T = ... = **1.103**

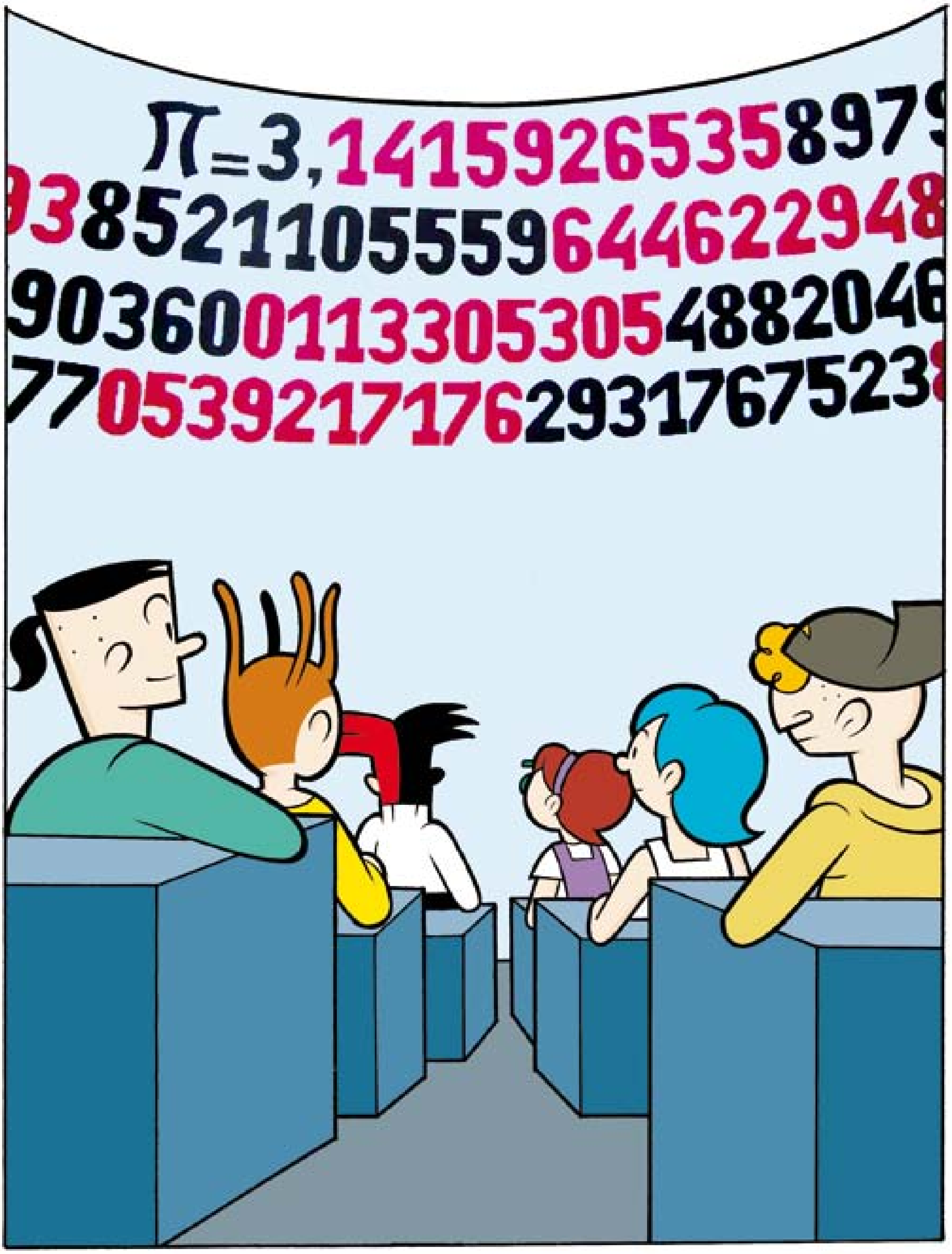
C = ... = **702**

Una vez que tengáis vuestros resultados efectuáis la siguiente operación:

$$\frac{2xT}{C}$$

Veréis que el número resultante se va aproximando estocásticamente a π a medida que las tiradas van siendo más numerosas.

Sala de π en el "Palais de la Decouverte". París



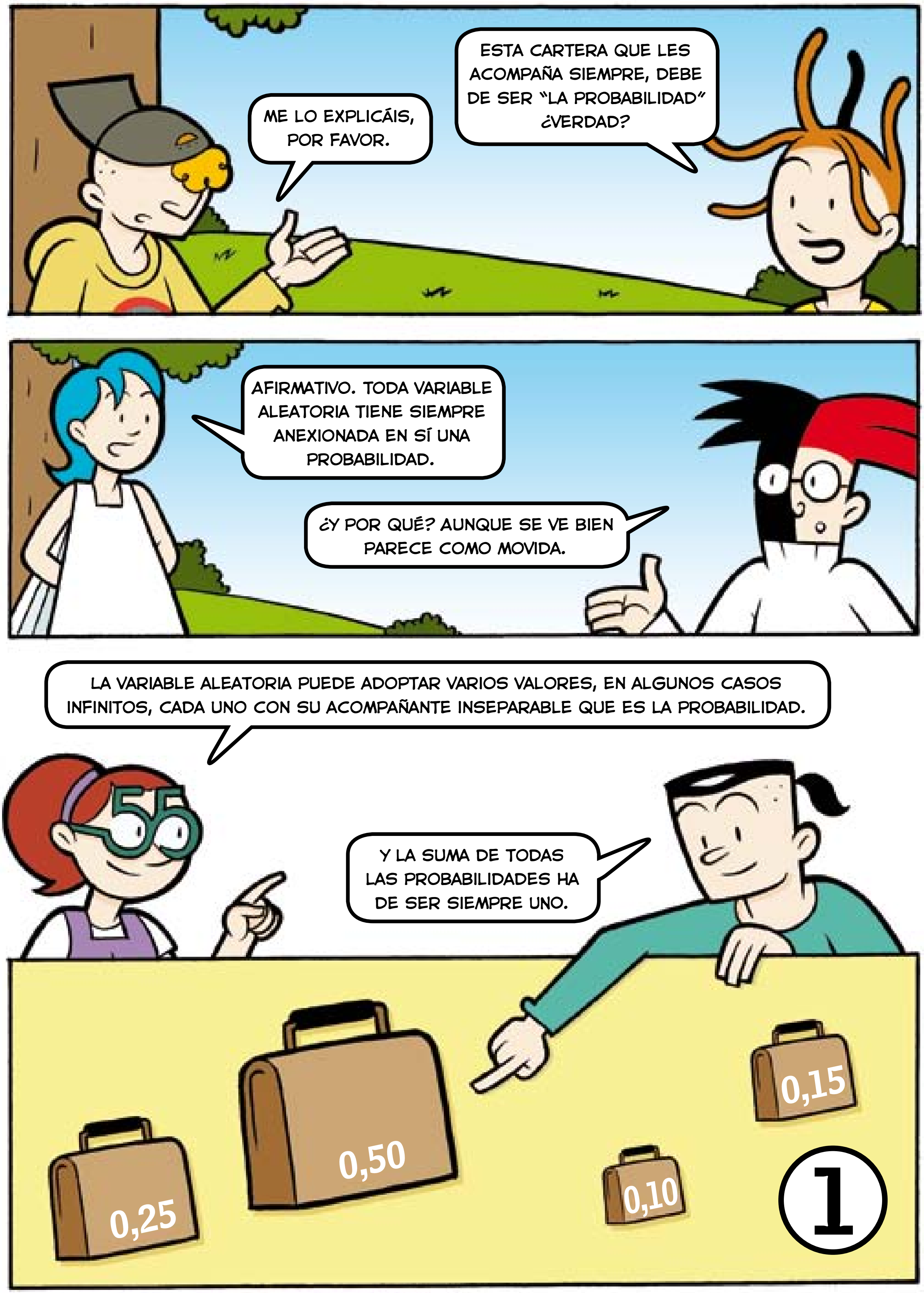
Capítulo 2

SIR FRANCIS GALTON
(1909)

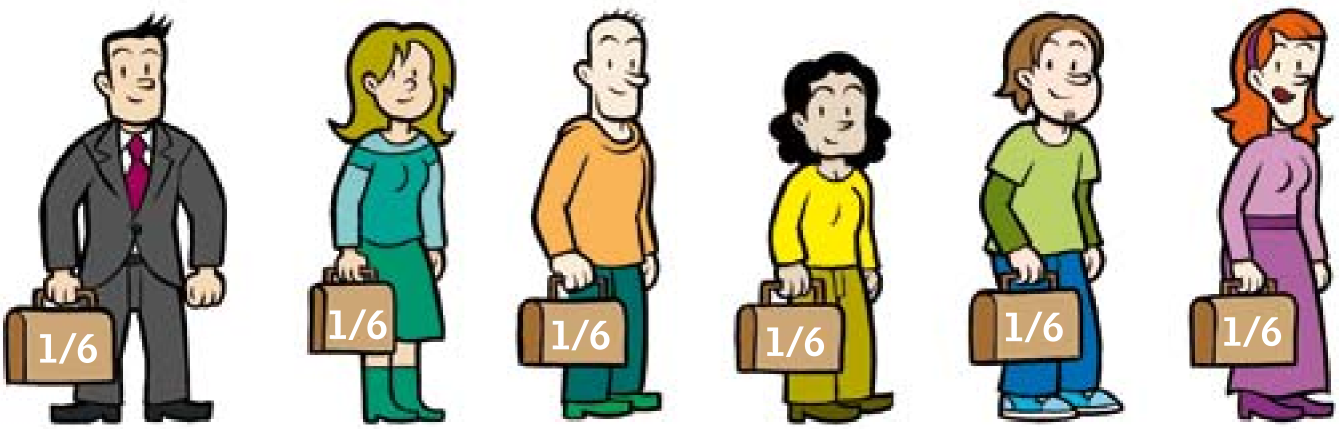
Duddeston, 1822 – Haslemere, 1911

Antropólogo y geógrafo, fue creador de la escuela biométrica inglesa.
Realizó, terminados sus estudios, grandes viajes con objetivo investigador, como su primo Charles Darwin.
Es uno de los introductores de los métodos estadísticos aplicados a la Biología conjuntamente con Karl Pearson, que curiosamente es autor de una biografía del personaje tratado.
Entre muchos otros temas fue el primero en explicar el fenómeno de regresión media y usar la línea de regresión; fue uno de los pioneros en el uso de la distribución normal.
Su ingenio le llevó a construir la máquina de Quincunx.

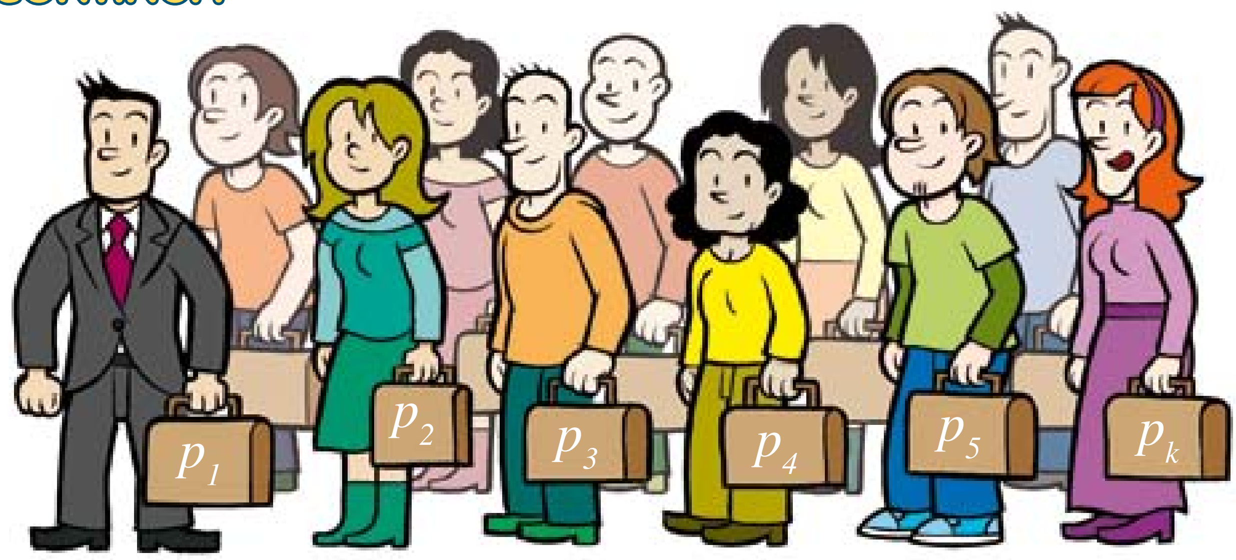

UN POCO DE AIRE NOS IRÁ BIEN PARA REFRESCAR LAS IDEAS. ¿OS ACORDÁIS DE CUANDO HABLÁBAMOS DE VARIABLES ALEATORIAS?
AQUELLO DEL SALTAMONTES Y EL CARACOL...
ESO ERA UNA CLASIFICACIÓN: EN DISCRETAS Y EN CONTINUAS.
ESPERAD, QUE ME GUSTARÍA HACER UN DIBUJO, POR SI... REFRESCA LA IDEA DE VARIABLE ALEATORIA.
PUES NO ESTÁ MAL. ¡QUEDAN MONOS!


ME LO EXPLICÁIS, POR FAVOR.
ESTA CARTERA QUE LES ACOMPAÑA SIEMPRE, DEBE DE SER "LA PROBABILIDAD" ¿VERDAD?
AFIRMATIVO. TODA VARIABLE ALEATORIA TIENE SIEMPRE ANEXIONADA EN SÍ UNA PROBABILIDAD.
¿Y POR QUÉ? AUNQUE SE VE BIEN PARECE COMO MOVIDA.
LA VARIABLE ALEATORIA PUEDE ADOPTAR VARIOS VALORES, EN ALGUNOS CASOS INFINITOS, CADA UNO CON SU ACOMPAÑANTE INSEPARABLE QUE ES LA PROBABILIDAD.
Y LA SUMA DE TODAS LAS PROBABILIDADES HA DE SER SIEMPRE UNO.
0,25
0,50
0,10
0,15
1

DISCRETA



CONTINUA



Sabiendo que $p_1+p_2+p_3+\ ...\ +p_k=1$

PONGAMOS ESTOS DOS DIBUJOS EN UN PLAN UN POCO MÁS SERIO.

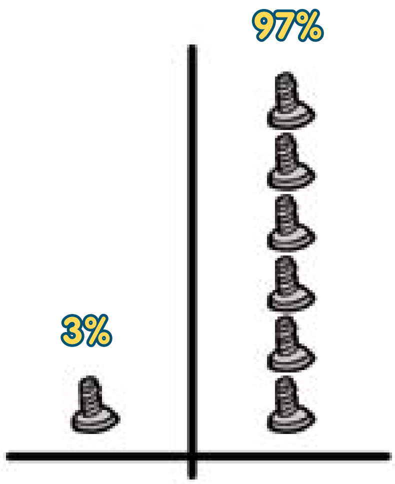
La tirada de un dado **DISCRETAS**

Igual probabilidad

Sabiendo que $p_1+p_2+p_3+ \ \dots + p_6 = 1 \ \Rightarrow \ p_i = \dfrac{1}{6}$

O SI LO MIRAMOS CON TORNILLOS, POR EJEMPLO.

DISCRETAS distinta probabilidad



Tornillos defectuosos 3%
Tornillos buenos 97%

Probabilidad $\dfrac{3}{100} + \dfrac{97}{100} = 1$


ESTO ES AQUELLO QUE DECÍA "SI LANZO UNA MONEDA AL AIRE, TENGO LA PROBABILIDAD 0'5 DE QUE SALGA CARA Y 0'5 DE QUE SALGA CRUZ".
QUE NO SERÁ LO MISMO QUE TENER MEDIA PROBABILIDAD DE CARA Y MEDIA DE CRUZ.
PERO... ¡QUÉ DICES!
LO VES, NO ACIERTO NI UNA...
55
MIRA, TÚ LO QUE TIENES QUE HACER ES...
PENSAR MUCHO LAS COSAS Y CUANDO LO HAYAS PENSADO BIEN...
DI LO CONTRARIO Y... ¡ACERTARÁS!
JA! JA! JA! JA! JA!
SI EL EXPERIMENTO DE BERNOULLI SE REPITE...
PODEMOS TENER LA DISTRIBUCIÓN BINOMIAL.
SÍ, PERO ANTES FIJAOS EN ESTO... UNA TRANSPARENCIA CON LAS CARACTERÍSTICAS MÁS IMPORTANTES.
55

DISTRIBUCIÓN DE BERNOULLI

$$P[X = x] = p^x \ (1 - p)^{1-x}$$

Siendo p la probabilidad de éxito y (1-p), como es lógico, la de fracaso.

$$P[X = x] \Rightarrow \begin{cases} 1 & \text{éxito} \\ 0 & \text{fracaso} \end{cases}$$

Su media o esperanza

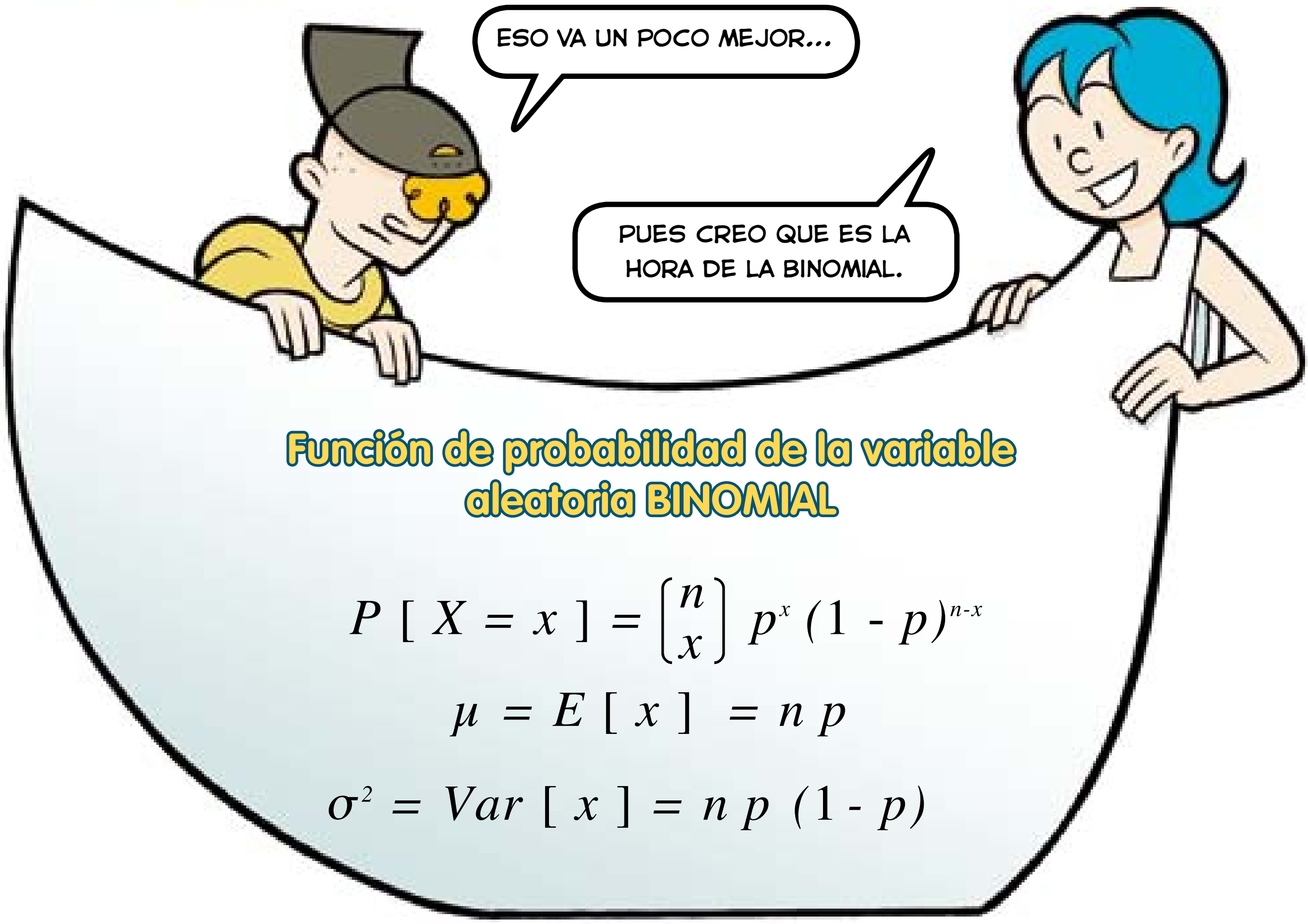
$$\mu = E[x] = p$$

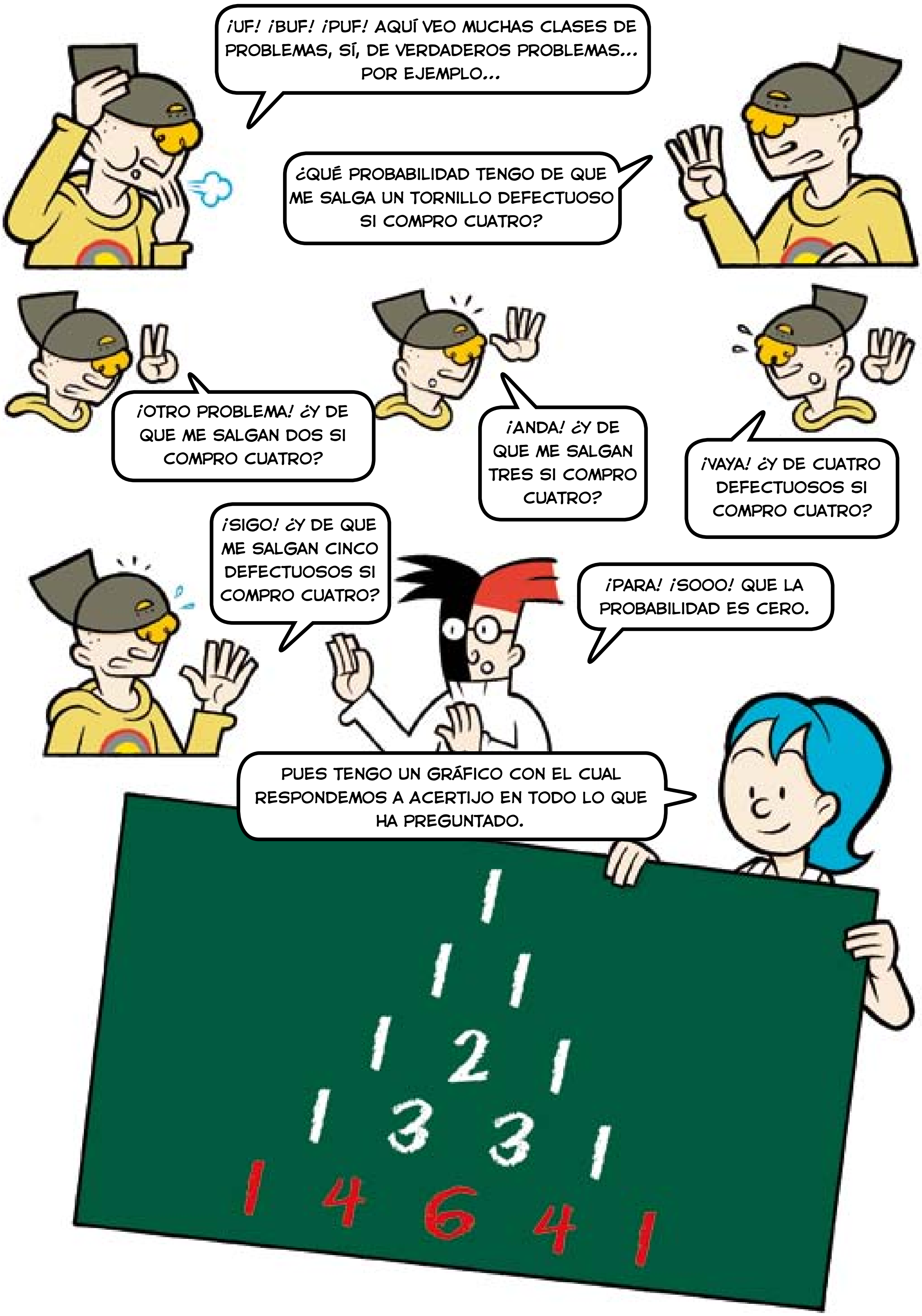
Su varianza

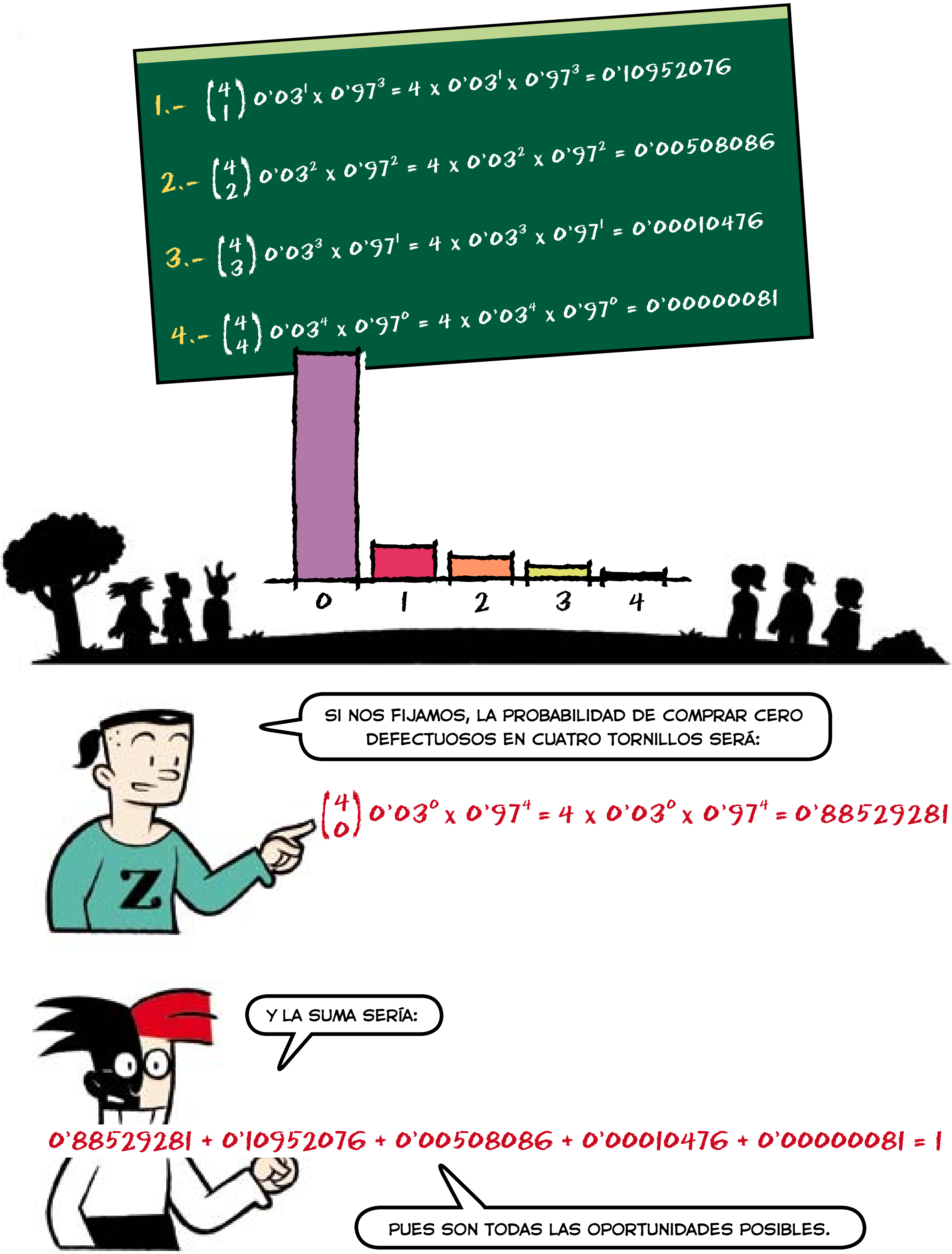
$$\sigma^2 = Var[x] = p(1 - p)$$


VA BASTANTE BIEN.
IR... TAL VEZ VA BASTANTE BIEN, PERO A MÍ, COMO NO ME DIGAS UN VERBIGRACIA, COMO DICE 55..., NO ME VIENE.

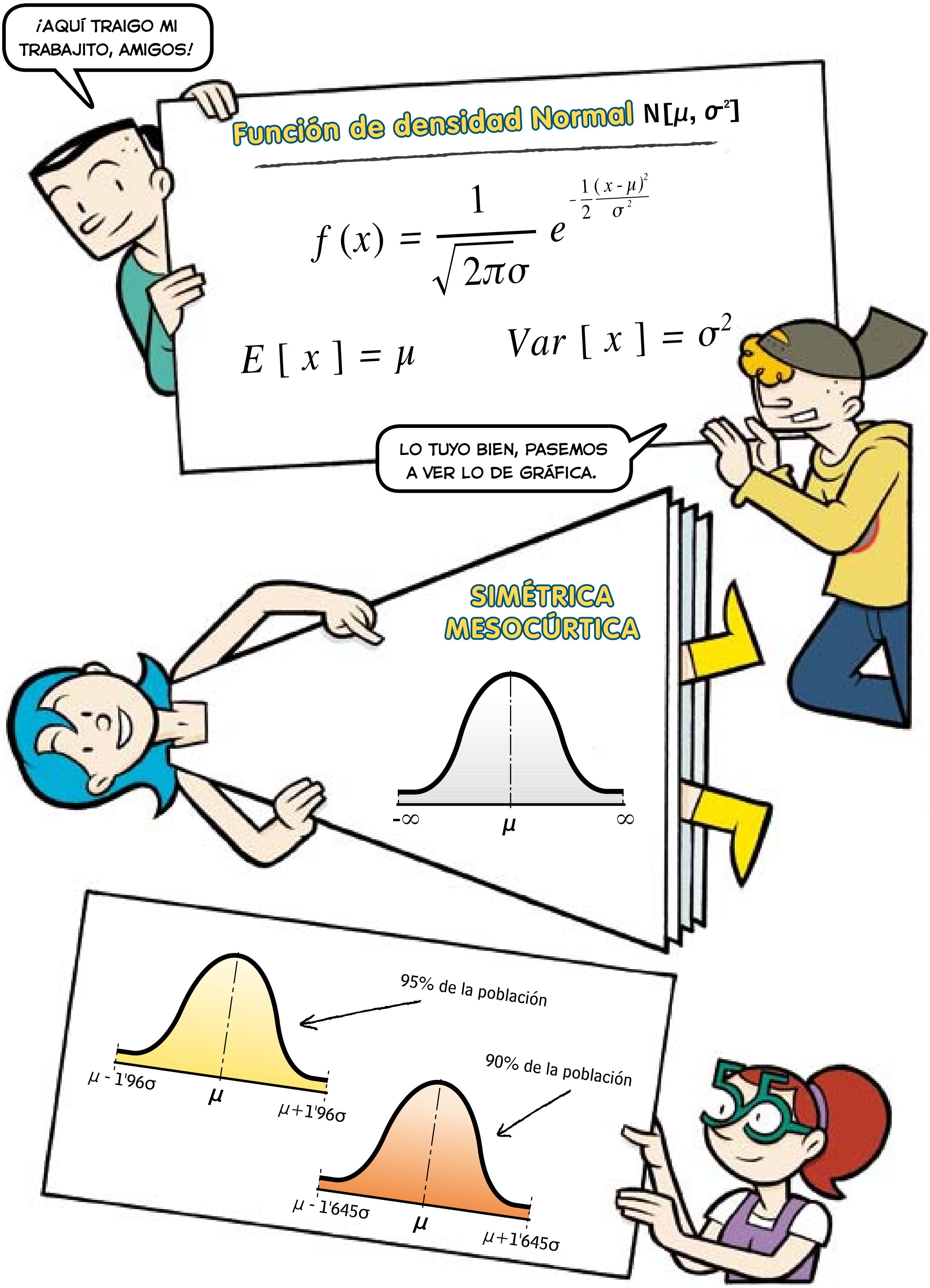

¿CUÁNTOS TORNILLOS DEFECTUOSOS ESPERAMOS HALLAR EN UNA CAJA DE 1.000 UNIDADES?
LA PROBABILIDAD DE SER defectuoso UN TORNILLO ES DE 0,03.
POR LO TANTO LA PROBABILIDAD DE no serlo ES DE 1 - 0,03 = 0,97.

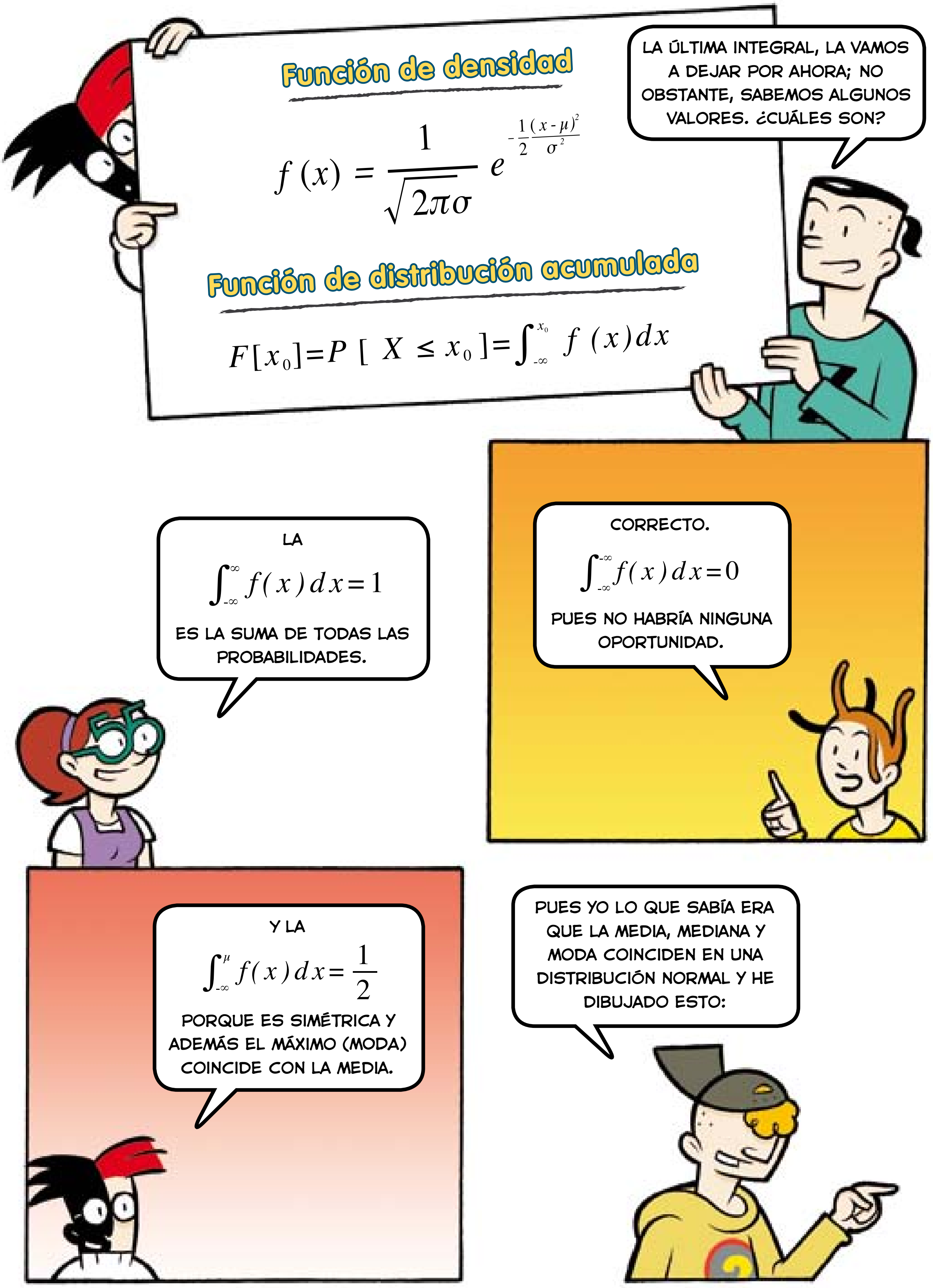

ESO VA UN POCO MEJOR...
PUES CREO QUE ES LA HORA DE LA BINOMIAL.
Función de probabilidad de la variable aleatoria BINOMIAL
$P[X = x] = \binom{n}{x} p^x (1 - p)^{n-x}$
$\mu = E[x] = n p$
$\sigma^2 = Var[x] = n p (1 - p)$


¡UF! ¡BUF! ¡PUF! AQUÍ VEO MUCHAS CLASES DE PROBLEMAS, SÍ, DE VERDADEROS PROBLEMAS... POR EJEMPLO...
¿QUÉ PROBABILIDAD TENGO DE QUE ME SALGA UN TORNILLO DEFECTUOSO SI COMPRO CUATRO?
¡OTRO PROBLEMA! ¿Y DE QUE ME SALGAN DOS SI COMPRO CUATRO?
¡ANDA! ¿Y DE QUE ME SALGAN TRES SI COMPRO CUATRO?
¡VAYA! ¿Y DE CUATRO DEFECTUOSOS SI COMPRO CUATRO?
¡SIGO! ¿Y DE QUE ME SALGAN CINCO DEFECTUOSOS SI COMPRO CUATRO?
¡PARA! ¡SOOO! QUE LA PROBABILIDAD ES CERO.
PUES TENGO UN GRÁFICO CON EL CUAL RESPONDEMOS A ACERTIJO EN TODO LO QUE HA PREGUNTADO.
1
1 1
1 2 1
1 3 3 1
1 4 6 4 1

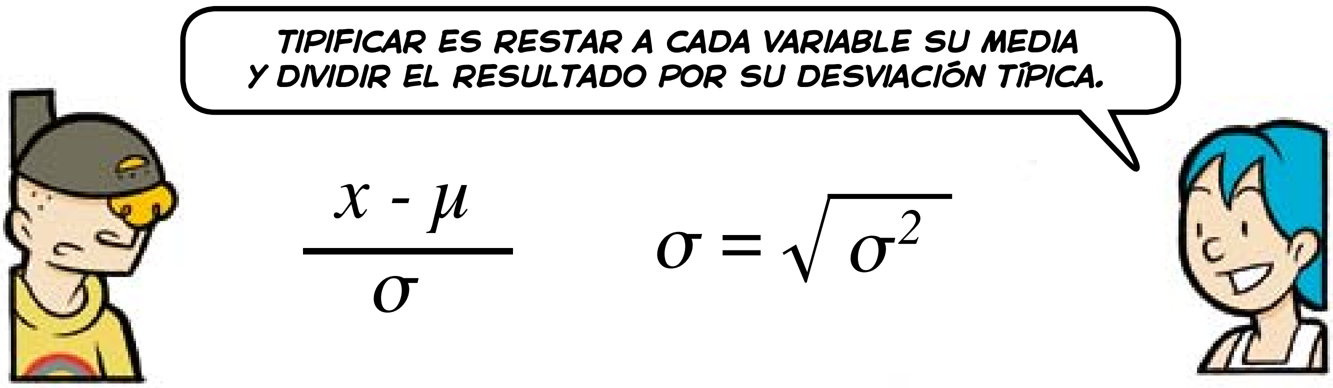

1.- $\binom{4}{1} 0'03^1 \times 0'97^3 = 4 \times 0'03^1 \times 0'97^3 = 0'10952076$
2.- $\binom{4}{2} 0'03^2 \times 0'97^2 = 4 \times 0'03^2 \times 0'97^2 = 0'00508086$
3.- $\binom{4}{3} 0'03^3 \times 0'97^1 = 4 \times 0'03^3 \times 0'97^1 = 0'00010476$
4.- $\binom{4}{4} 0'03^4 \times 0'97^0 = 4 \times 0'03^4 \times 0'97^0 = 0'00000081$
0
1
2
3
4
SI NOS FIJAMOS, LA PROBABILIDAD DE COMPRAR CERO DEFECTUOSOS EN CUATRO TORNILLOS SERÁ:
$\binom{4}{0} 0'03^0 \times 0'97^4 = 4 \times 0'03^0 \times 0'97^4 = 0'88529281$
Y LA SUMA SERÍA:
$0'88529281 + 0'10952076 + 0'00508086 + 0'00010476 + 0'00000081 = 1$
PUES SON TODAS LAS OPORTUNIDADES POSIBLES.


SÍ, PERO DE MI PREGUNTA, ¿QUÉ?
LOS TORNILLOS DEFECTUOSOS QUE "**ESPERAMOS**" AL COMPRAR 1.000.
LA ESPERANZA, O SEA, LA MEDIA...
E (x) = np = 1000 x 0'03 = 30 TORNILLOS DEFECTUOSOS.
YA ESTOY AGOTADO, ¿POR QUÉ NO LO DEJAMOS PARA LA SEMANA QUE VIENE?... Y A VER SI DAMOS EL CAMPANAZO DE GAUSS HABLANDO DE LA DISTRIBUCIÓN NORMAL.
¡VALE!

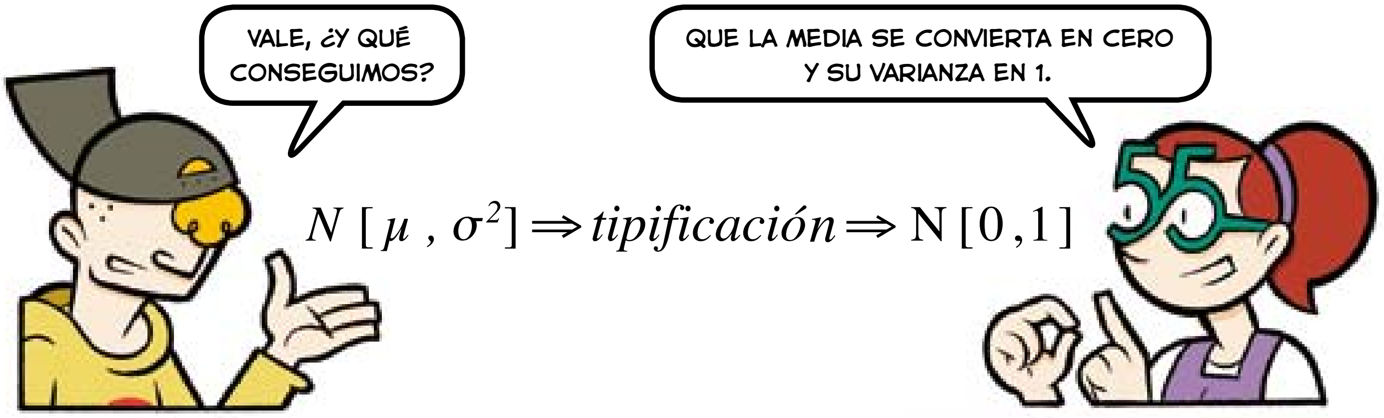

¡AQUÍ TRAIGO MI TRABAJITO, AMIGOS!
Función de densidad Normal N[μ, σ²]
$f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}$
$E[x] = \mu$
$Var[x] = \sigma^2$
LO TUYO BIEN, PASEMOS A VER LO DE GRÁFICA.
SIMÉTRICA
MESOCÚRTICA
-∞
μ
∞
95% de la población
μ - 1'96σ
μ
μ+1'96σ
90% de la población
μ - 1'645σ
μ
μ+1'645σ



Función de densidad
$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}$$
Función de distribución acumulada
$$F[x_0] = P[X \leq x_0] = \int_{-\infty}^{x_0} f(x)dx$$
LA ÚLTIMA INTEGRAL, LA VAMOS A DEJAR POR AHORA; NO OBSTANTE, SABEMOS ALGUNOS VALORES. ¿CUÁLES SON?
LA $\int_{-\infty}^{\infty} f(x)dx = 1$ ES LA SUMA DE TODAS LAS PROBABILIDADES.
CORRECTO. $\int_{-\infty}^{-\infty} f(x)dx = 0$ PUES NO HABRÍA NINGUNA OPORTUNIDAD.
Y LA $\int_{-\infty}^{\mu} f(x)dx = \frac{1}{2}$ PORQUE ES SIMÉTRICA Y ADEMÁS EL MÁXIMO (MODA) COINCIDE CON LA MEDIA.
PUES YO LO QUE SABÍA ERA QUE LA MEDIA, MEDIANA Y MODA COINCIDEN EN UNA DISTRIBUCIÓN NORMAL Y HE DIBUJADO ESTO:



$$\frac{x - \mu}{\sigma} \qquad \sigma = \sqrt{\sigma^2}$$



N[0,1]

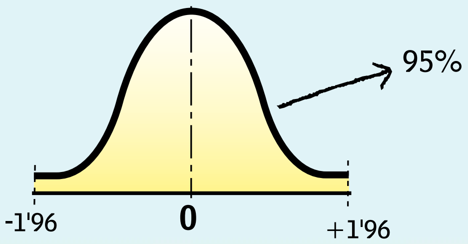
Función de densidad

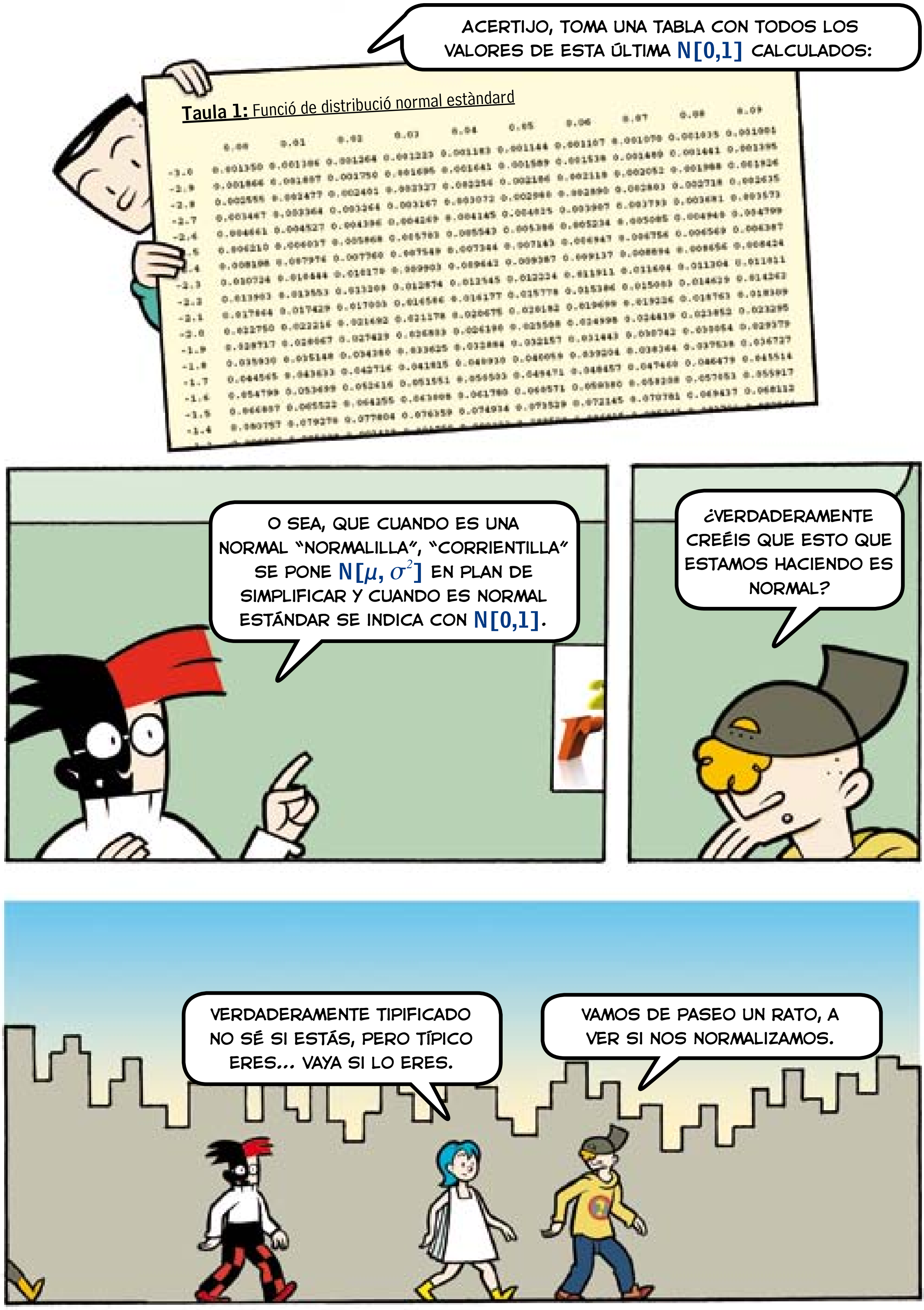
$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

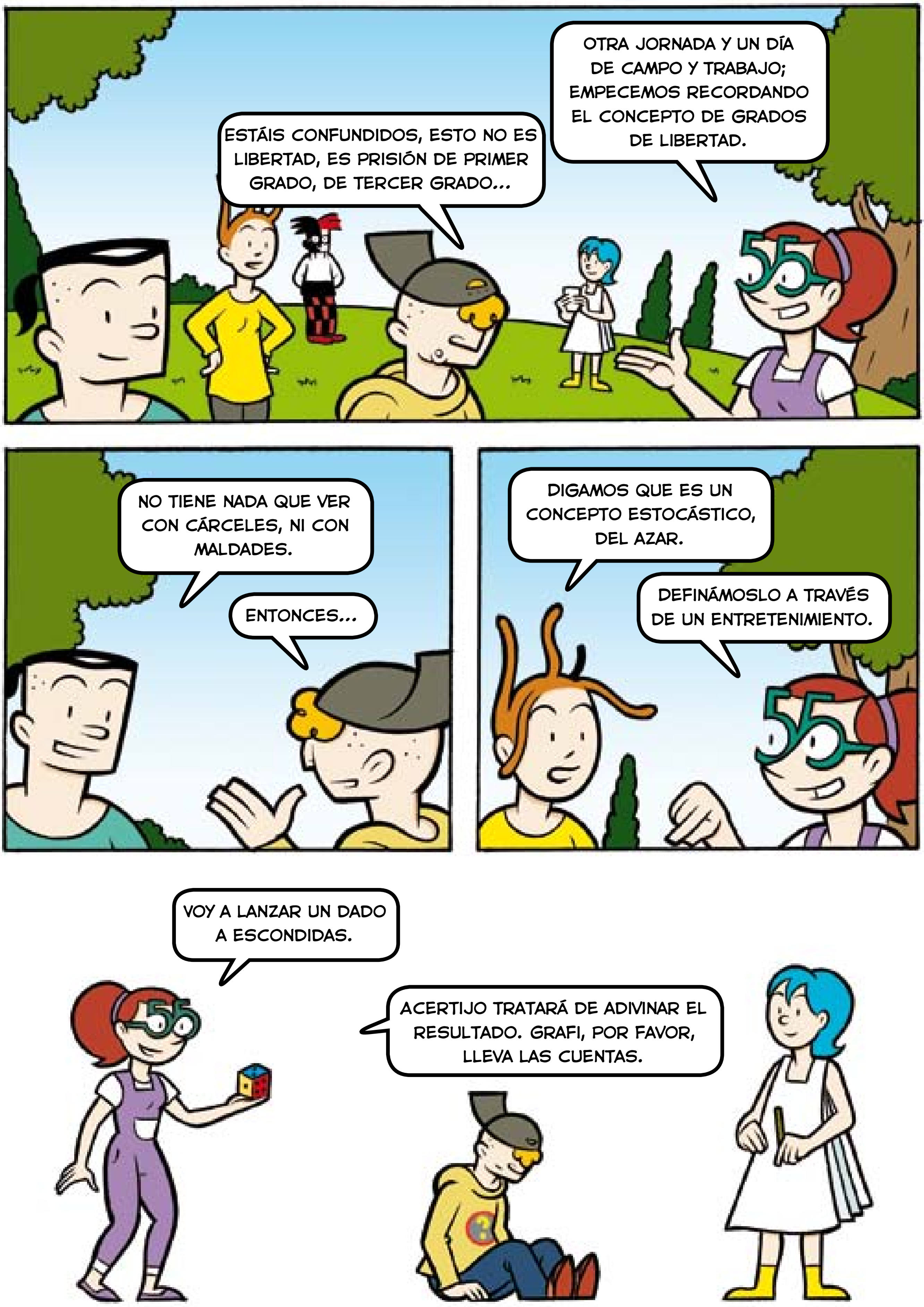
Función de distribución acumulada

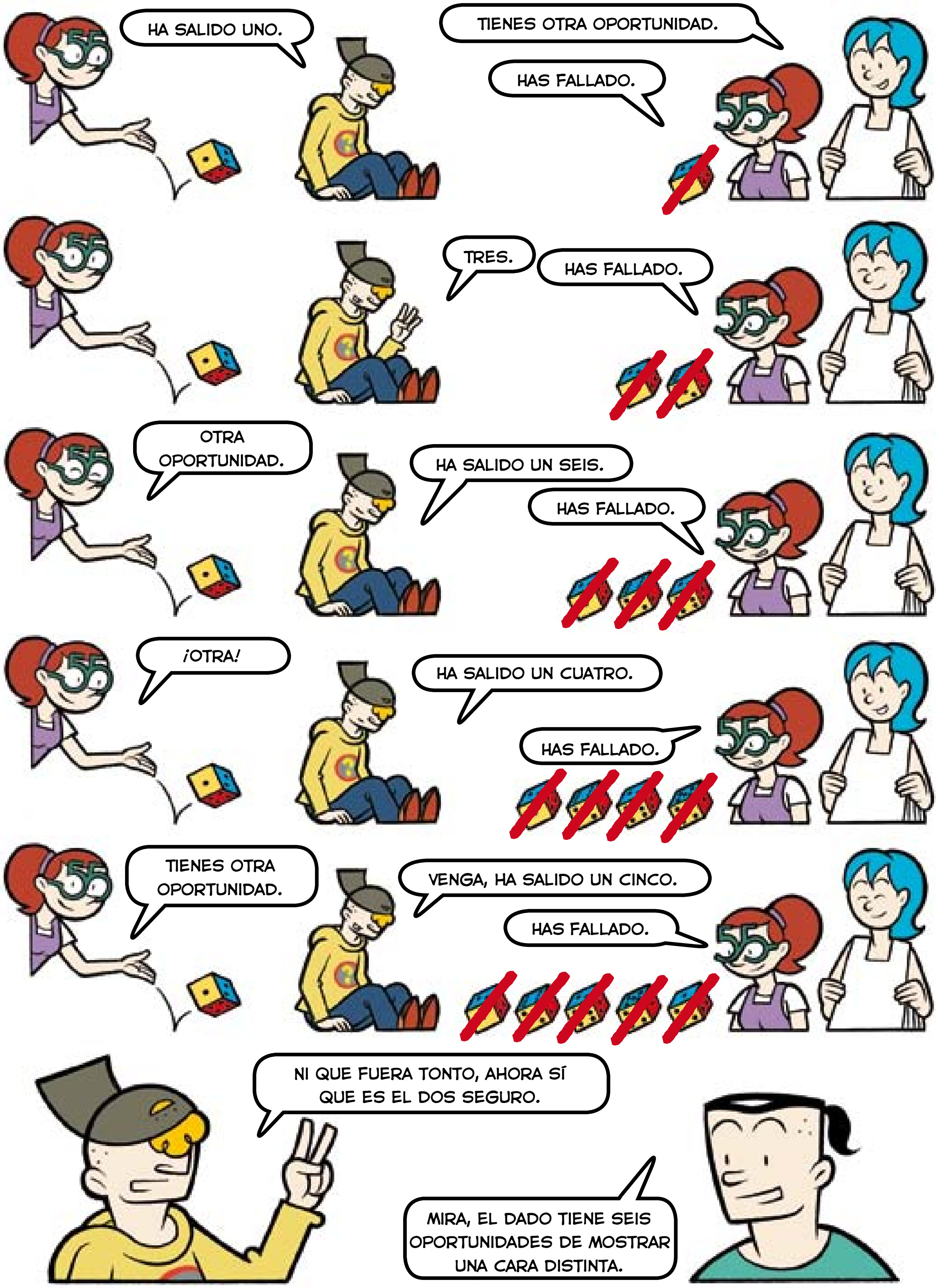
$$F(x_0) = \int_{-\infty}^{x_0} f(x)dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x_0} e^{-\frac{x^2}{2}} dx$$

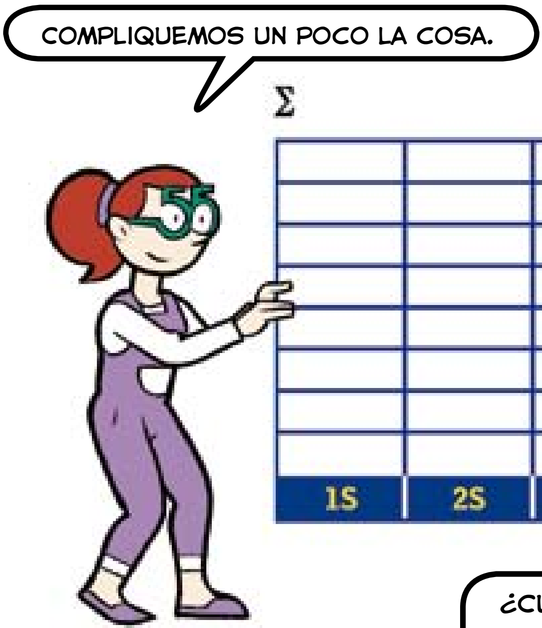
$$E[x] = 0 \qquad \sigma^2 = 1 \qquad \sigma = 1$$




ACERTIJO, TOMA UNA TABLA CON TODOS LOS VALORES DE ESTA ÚLTIMA N[0,1] CALCULADOS:
Taula 1: Funció de distribució normal estàndard
O SEA, QUE CUANDO ES UNA NORMAL "NORMALILLA", "CORRIENTILLA" SE PONE $N[\mu, \sigma^2]$ EN PLAN DE SIMPLIFICAR Y CUANDO ES NORMAL ESTÁNDAR SE INDICA CON N[0,1].
¿VERDADERAMENTE CREÉIS QUE ESTO QUE ESTAMOS HACIENDO ES NORMAL?
VERDADERAMENTE TIPIFICADO NO SÉ SI ESTÁS, PERO TÍPICO ERES... VAYA SI LO ERES.
VAMOS DE PASEO UN RATO, A VER SI NOS NORMALIZAMOS.


OTRA JORNADA Y UN DÍA DE CAMPO Y TRABAJO; EMPECEMOS RECORDANDO EL CONCEPTO DE GRADOS DE LIBERTAD.
ESTÁIS CONFUNDIDOS, ESTO NO ES LIBERTAD, ES PRISIÓN DE PRIMER GRADO, DE TERCER GRADO...
NO TIENE NADA QUE VER CON CÁRCELES, NI CON MALDADES.
ENTONCES...
DIGAMOS QUE ES UN CONCEPTO ESTOCÁSTICO, DEL AZAR.
DEFINÁMOSLO A TRAVÉS DE UN ENTRETENIMIENTO.
VOY A LANZAR UN DADO A ESCONDIDAS.
ACERTIJO TRATARÁ DE ADIVINAR EL RESULTADO. GRAFI, POR FAVOR, LLEVA LAS CUENTAS.


HA SALIDO UNO.
TIENES OTRA OPORTUNIDAD.
HAS FALLADO.
TRES.
HAS FALLADO.
OTRA OPORTUNIDAD.
HA SALIDO UN SEIS.
HAS FALLADO.
¡OTRA!
HA SALIDO UN CUATRO.
HAS FALLADO.
TIENES OTRA OPORTUNIDAD.
VENGA, HA SALIDO UN CINCO.
HAS FALLADO.
NI QUE FUERA TONTO, AHORA SÍ QUE ES EL DOS SEGURO.
MIRA, EL DADO TIENE SEIS OPORTUNIDADES DE MOSTRAR UNA CARA DISTINTA.



Σ

					S1
					S2
					S3
					S4
					S5
					S6
					S7
					S8
1S	2S	3S	4S	5S	

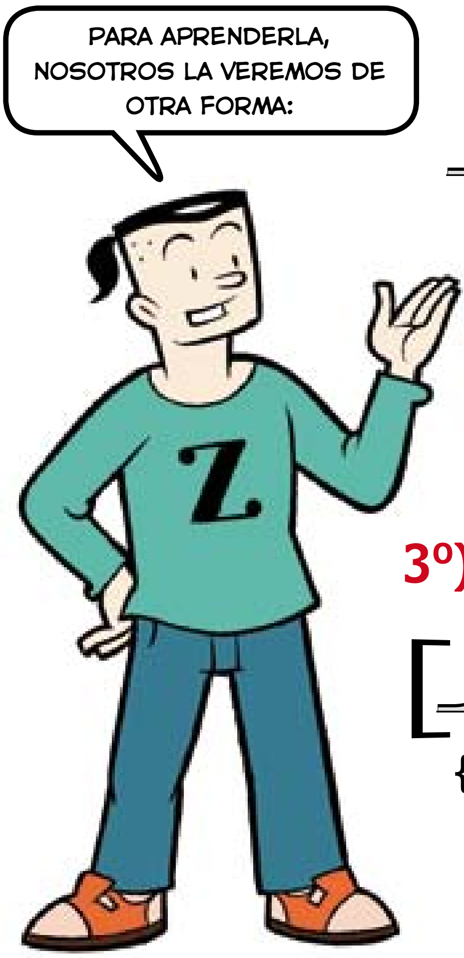
HAY QUE RELLENAR LAS CASILLAS CON NÚMEROS CUALESQUIERA, CON LA CONDICIÓN DE QUE SUMEN LOS TOTALES YA DEFINIDOS, TANTO EN FILAS COMO EN COLUMNAS. (LOS NÚMEROS PUEDEN SER NEGATIVOS).

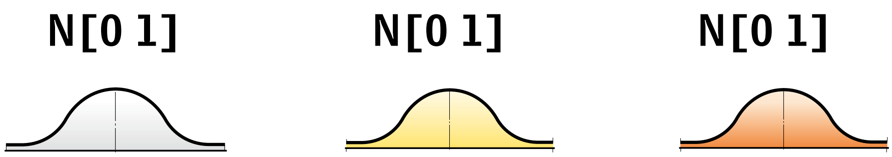
2	12	3	11		40
10					S2
7					S3
5					S4
5					S5
18					S6
4					S7
					S8
52	2S	3S	4S	5S	

4 x 7 = 28 grados de libertad.

[NÚMERO COLUMNAS – 1] x [NÚMERO FILAS – 1]

1º) Tomamos un cierto número de distribuciones normales estándar.

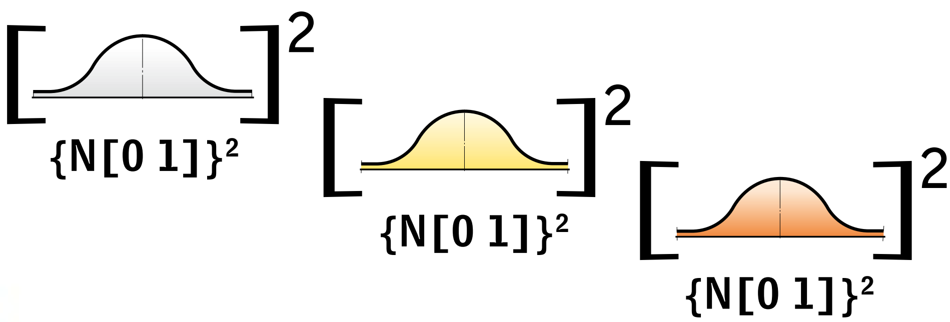




2º) Elegiremos las que sean independientes entre sí.

a) Puede ser sólo una.
b) Podrían ser de la forma N (μ 1) pero nosotros supondremos siempre que son estándar, por simplificar su estudio.

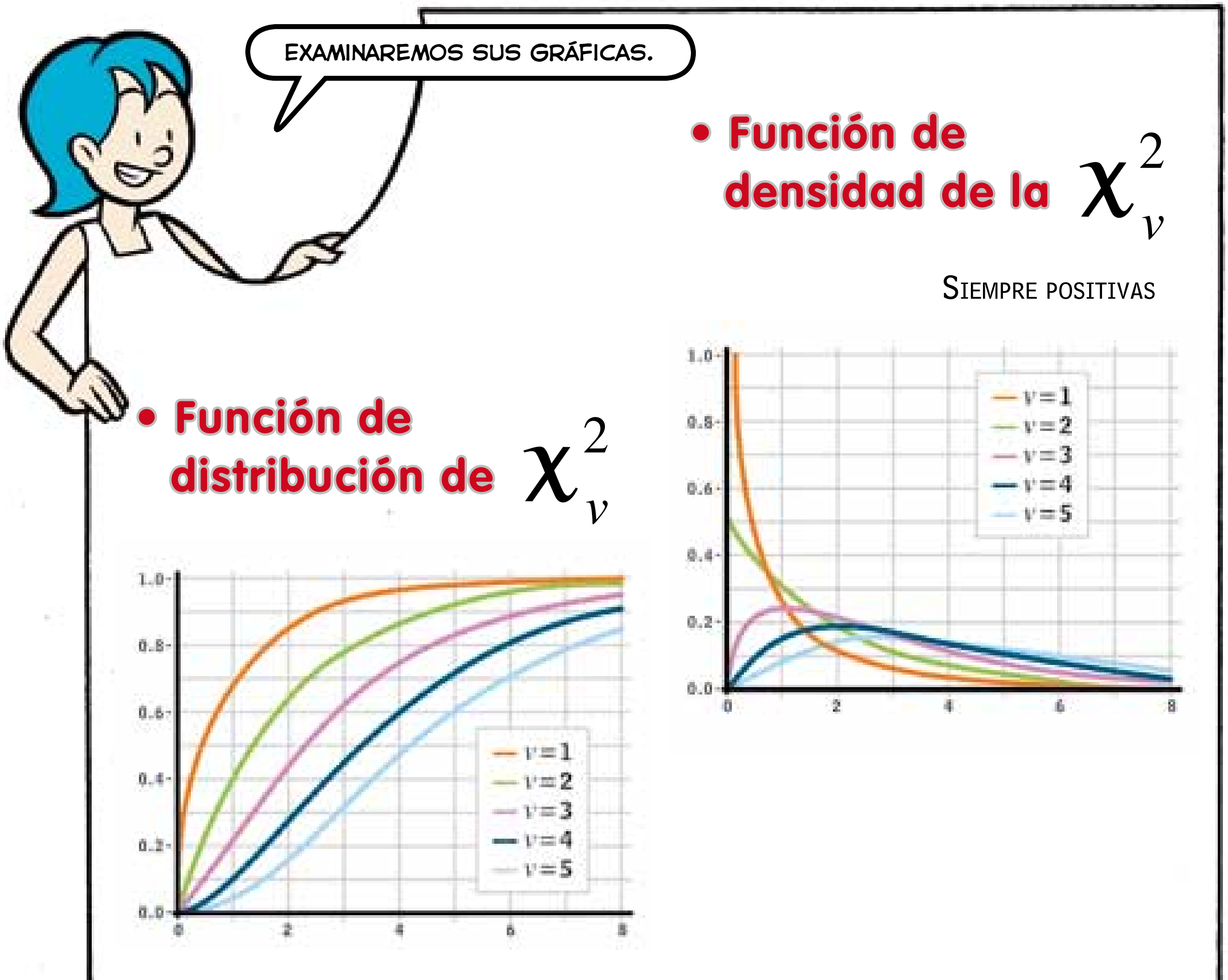
3º) Las elevamos al cuadrado:



4º) **Las sumamos.**

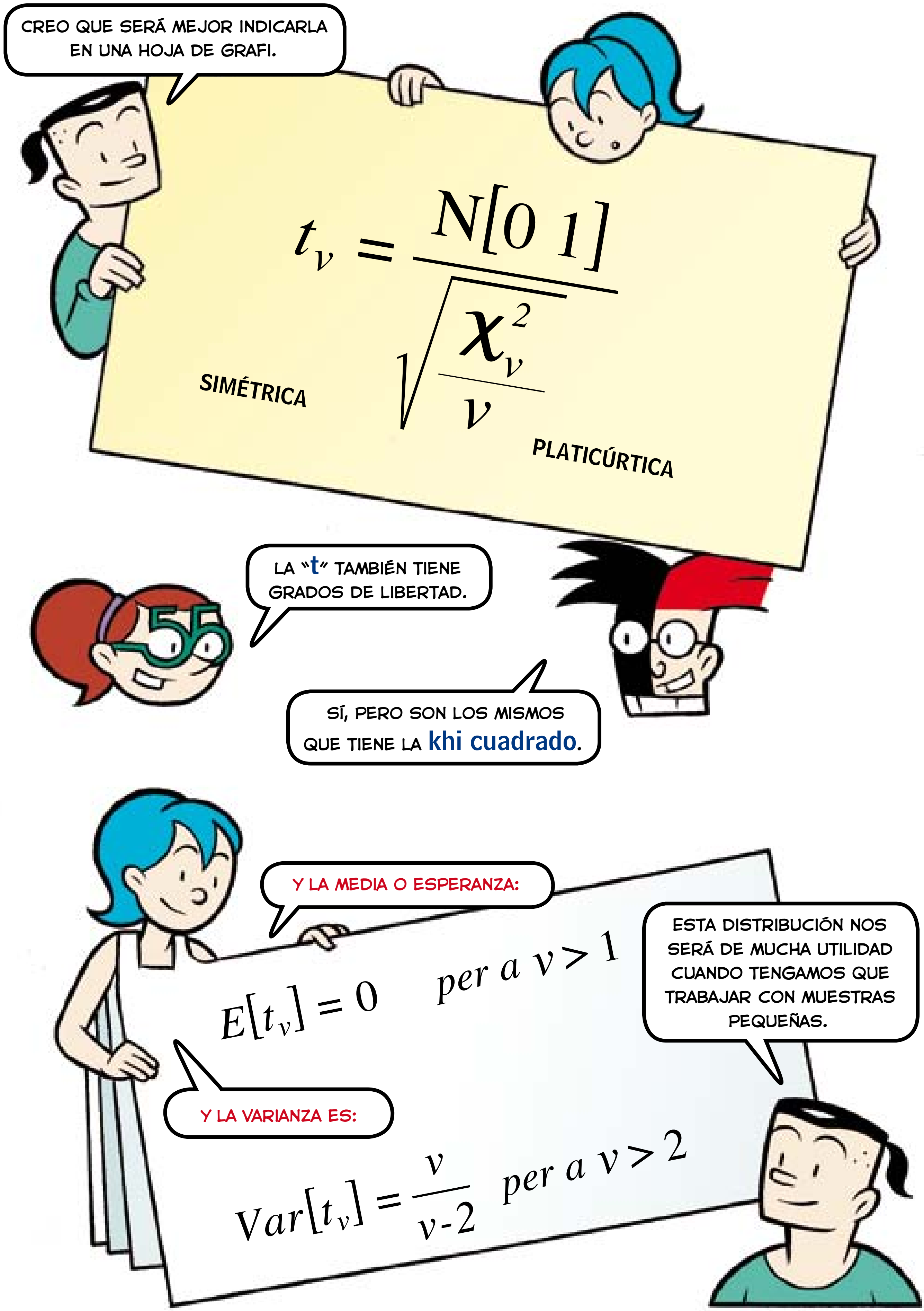
$$\sum_{i=1}^{v}[NID_i(0\,1)]^2 = \chi_v^2$$

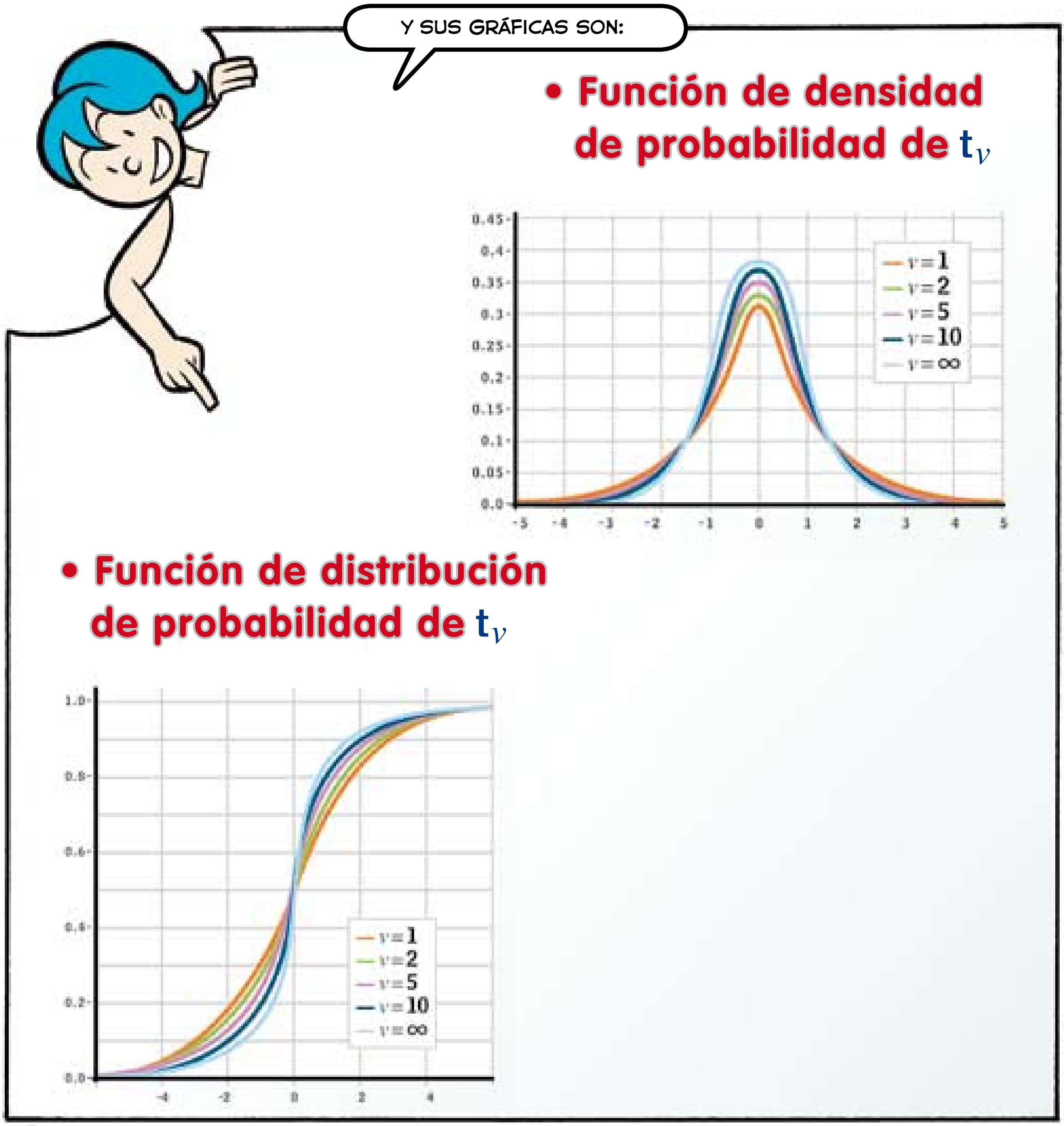
NID: Normales independientemente distribuidas.



Las representaciones son distintas según los grados de libertad.

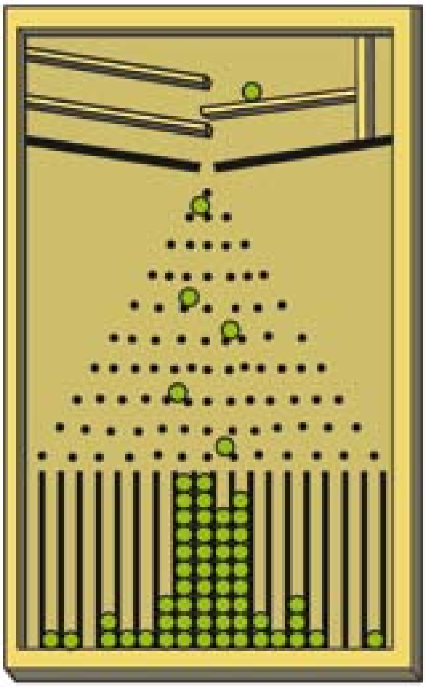

PUES LO ÚNICO QUE ME QUEDA POR DECIR ES:
¿QUÉ?
QUE ES PRODUCTIVA.
QUIERO DECIR... LO SIGUIENTE:
¡UF! VALE...
$X_5^2 + X_7^2 = X_{12}^2$
$X_{21}^2 + X_1^2 - X_3^2 = X_{19}^2$
ESTO SE ANIMA. ACORDAOS DE AQUEL SEÑOR... ¿CÓMO SE LLAMABA? SEALEY GOSSET, QUE PARA PUBLICAR SUS ESTUDIOS TENÍA QUE PONER UN SEUDÓNIMO...
STUDENT.
¡SÍ! PUES TIENE UNA DISTRIBUCIÓN, LA "t" DE STUDENT.
POBRE HOMBRE, CON SEUDÓNIMO Y SU DISTRIBUCIÓN CON MINÚSCULA.


CREO QUE SERÁ MEJOR INDICARLA EN UNA HOJA DE GRAFI.
$t_\nu = \frac{N[0\ 1]}{\sqrt{\frac{\chi^2_\nu}{\nu}}}$
SIMÉTRICA
PLATICÚRTICA
LA "t" TAMBIÉN TIENE GRADOS DE LIBERTAD.
SÍ, PERO SON LOS MISMOS QUE TIENE LA khi cuadrado.
Y LA MEDIA O ESPERANZA:
$E[t_\nu] = 0 \quad per\ a\ \nu > 1$
Y LA VARIANZA ES:
$Var[t_\nu] = \frac{\nu}{\nu\text{-}2} \quad per\ a\ \nu > 2$
ESTA DISTRIBUCIÓN NOS SERÁ DE MUCHA UTILIDAD CUANDO TENGAMOS QUE TRABAJAR CON MUESTRAS PEQUEÑAS.


Y SUS GRÁFICAS SON:
• Función de densidad de probabilidad de t_ν
$\nu=1$
$\nu=2$
$\nu=5$
$\nu=10$
$\nu=\infty$
• Función de distribución de probabilidad de t_ν
$\nu=1$
$\nu=2$
$\nu=5$
$\nu=10$
$\nu=\infty$


UN PEQUEÑO DETALLE: $t_\infty \approx \mathrm{N}[0\,1]$
¡UN GRAN DETALLE!
¡YA ESTÁ BIEEEEN!

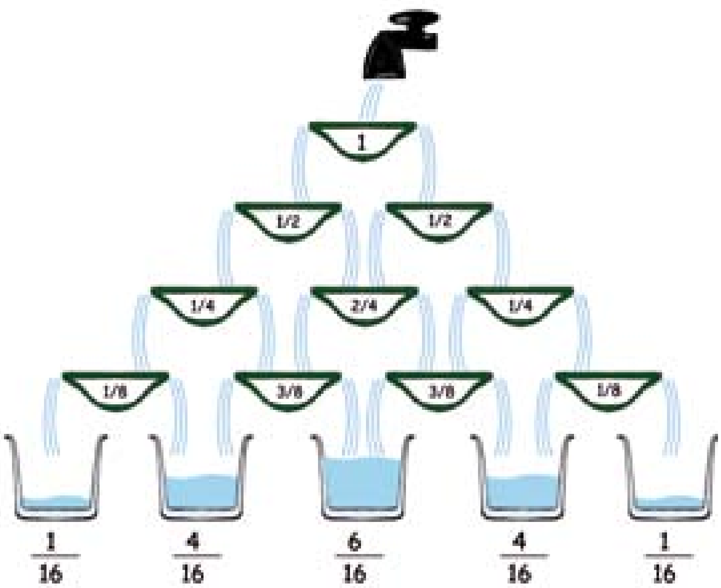
Máquina de Galton



Caja con artilugio para que vayan cayendo las bolitas, que irán chocando aleatoriamente con unos clavos situados como se ve en la figura, para al final caer sobre los casilleros de la base.

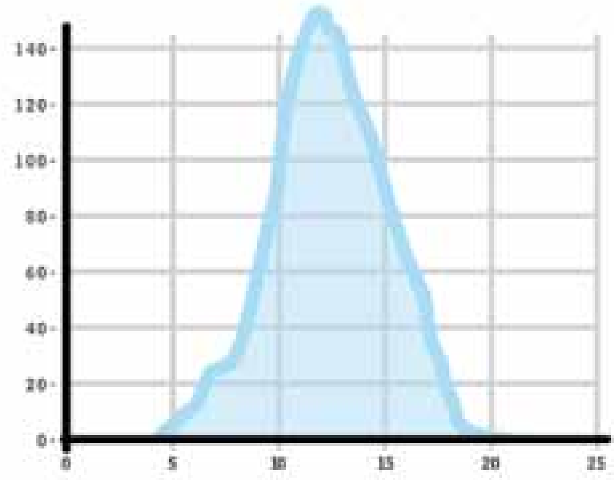
Las bolitas, al chocar con los clavos, tienen la misma probabilidad de ir a la derecha que a la izquierda.

Recordemos una experiencia realizada en el primer tomo.

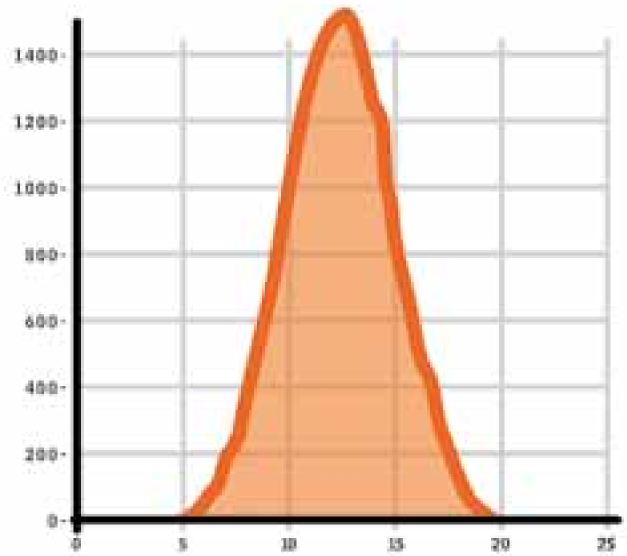


Tendremos como resultado, cuando el número de bolitas sea lo suficientemente grande, una colocación sobre los casilleros que se acercará a la forma de la distribución normal.

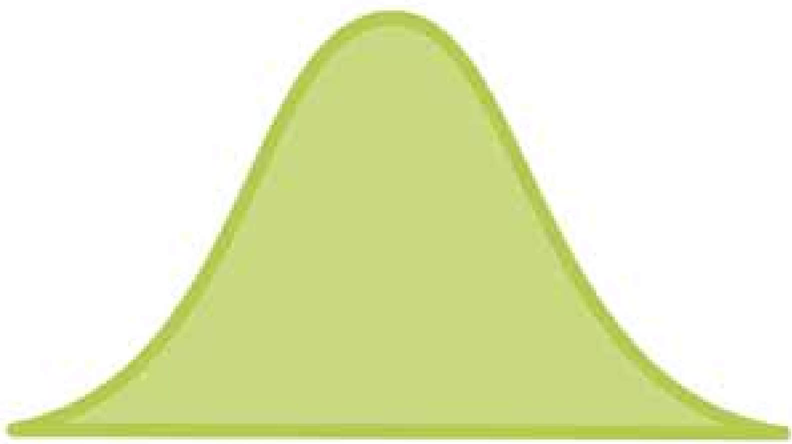
1.000 bolitas



100.000 bolitas



infinitas bolitas



Capítulo 3

PAFNUTI CHEBYSHEV

1821 - 1894

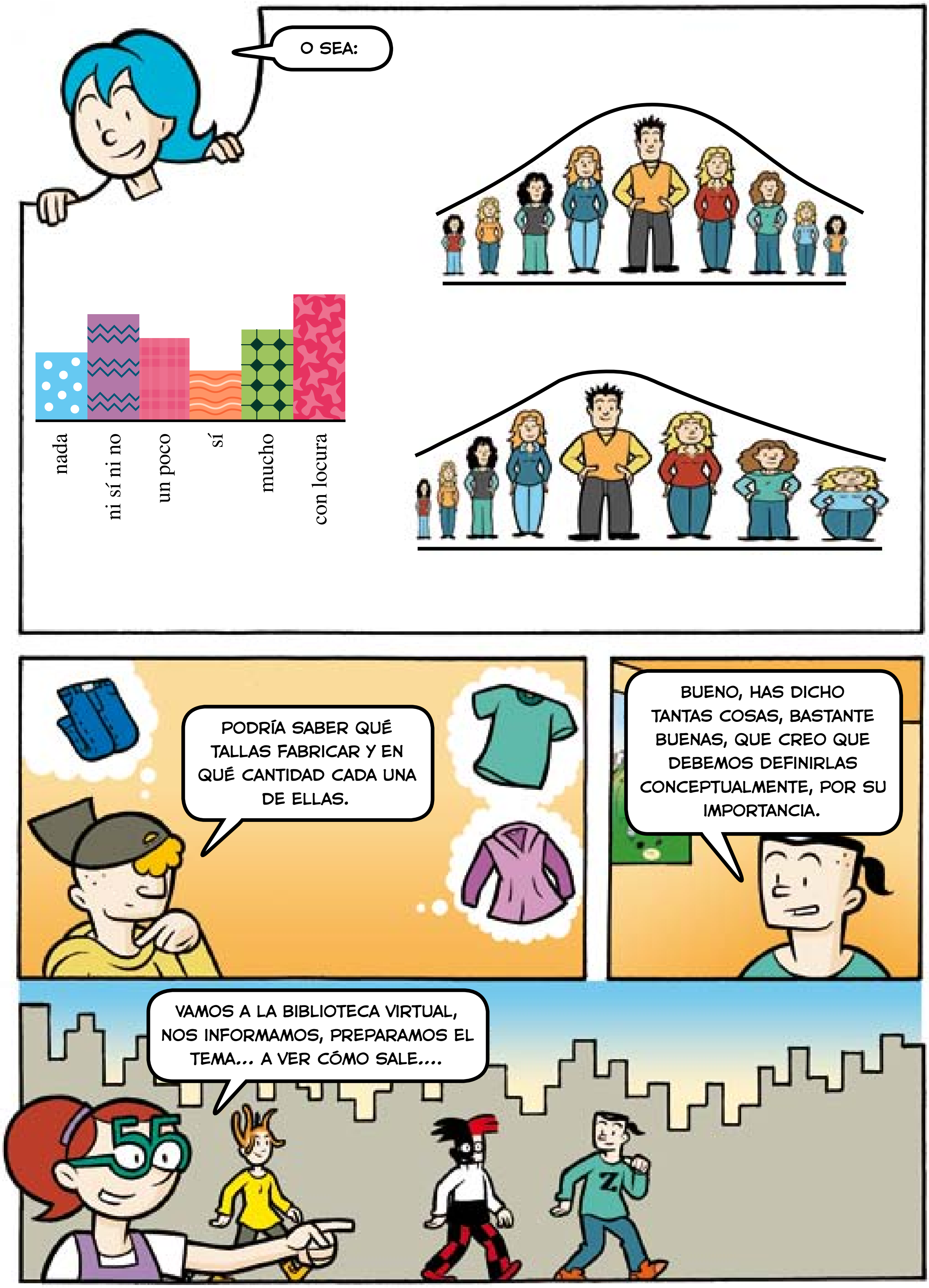
Matemático ruso también conocido como Tchebychew, Chebychev o Cebisev.
Defendió la tesis *Un intento de análisis elemental de la teoría probabilística*.
Recibió la medalla de plata por el trabajo *Cálculo de las raíces de las ecuaciones*.
La disertación sobre la integración con ayuda de algoritmos le hizo conseguir la plaza de Profesor Titular en la Universidad de San Petersburgo.
Pero es más conocido entre los estudiantes de estadística, principalmente, por su Desigualdad de Chebyshov.

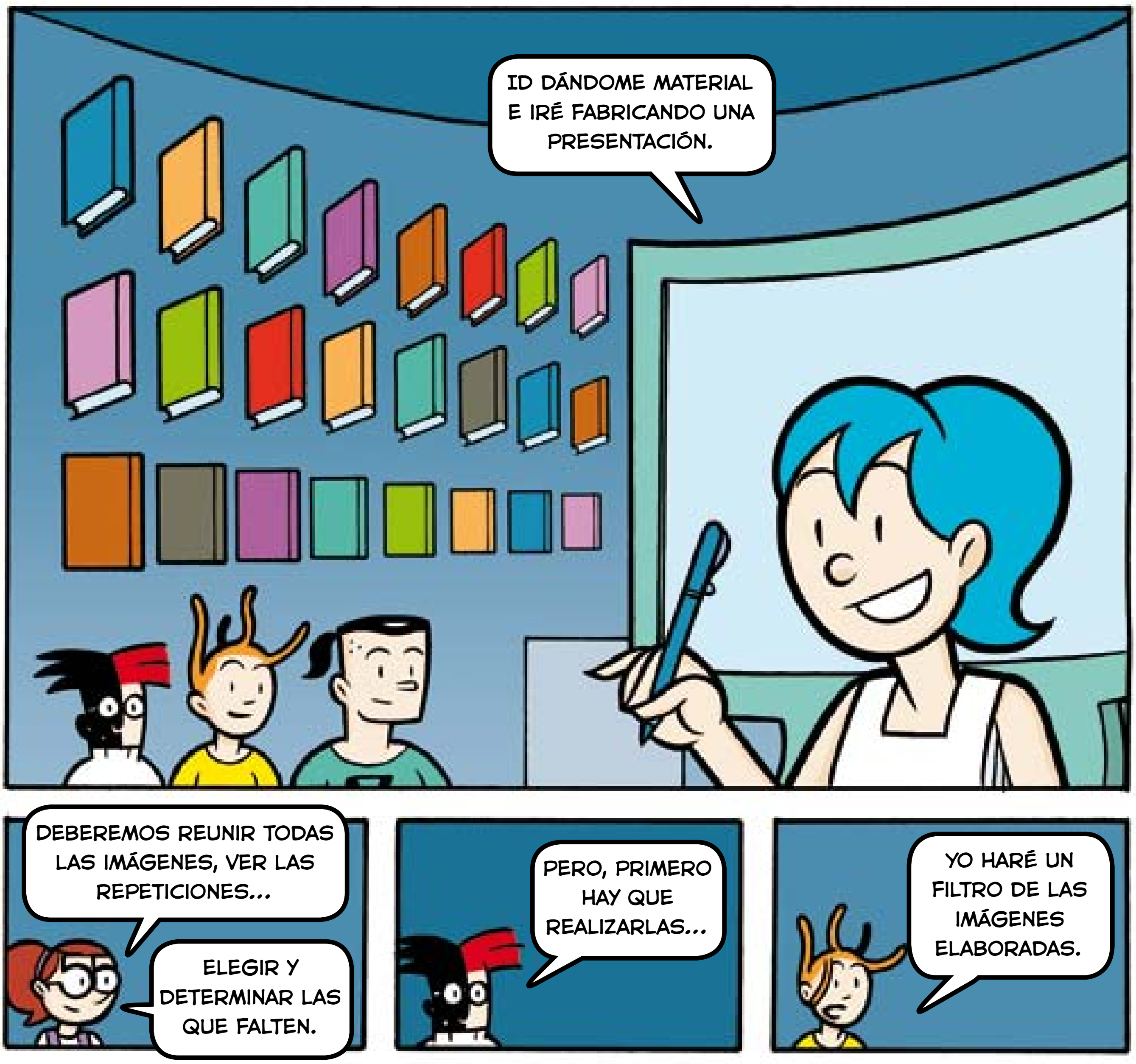
Que en un caso práctico sería:

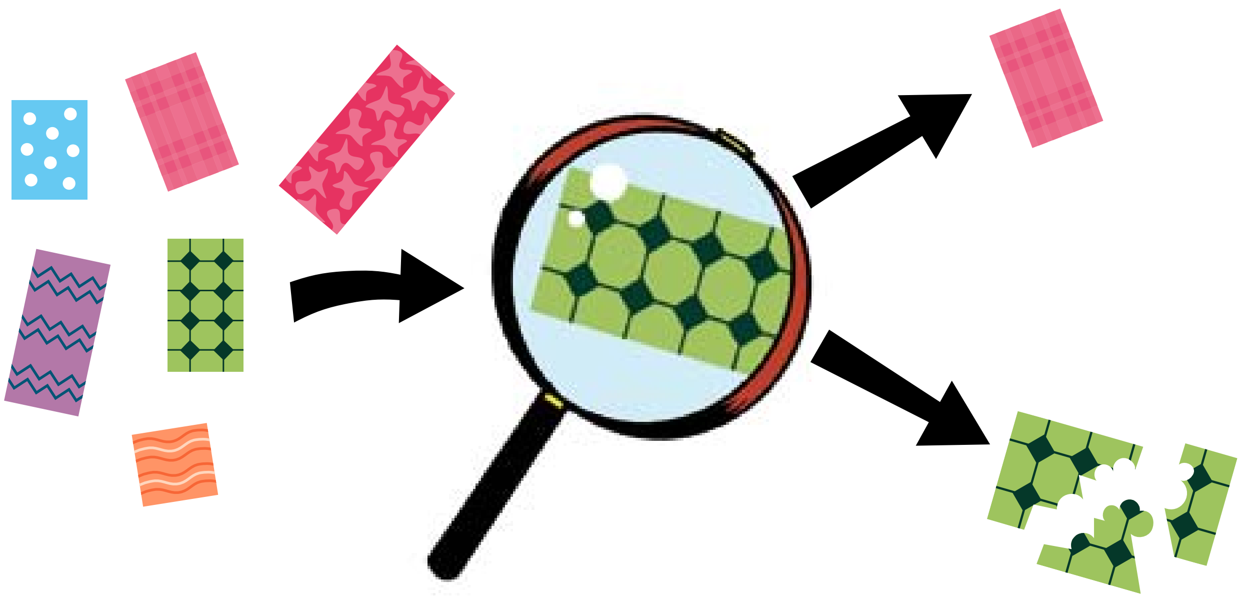
$$P\left[\left|x - E\left[x\right]\right| \geq k\sigma\right] \leq \frac{1}{k^2}$$

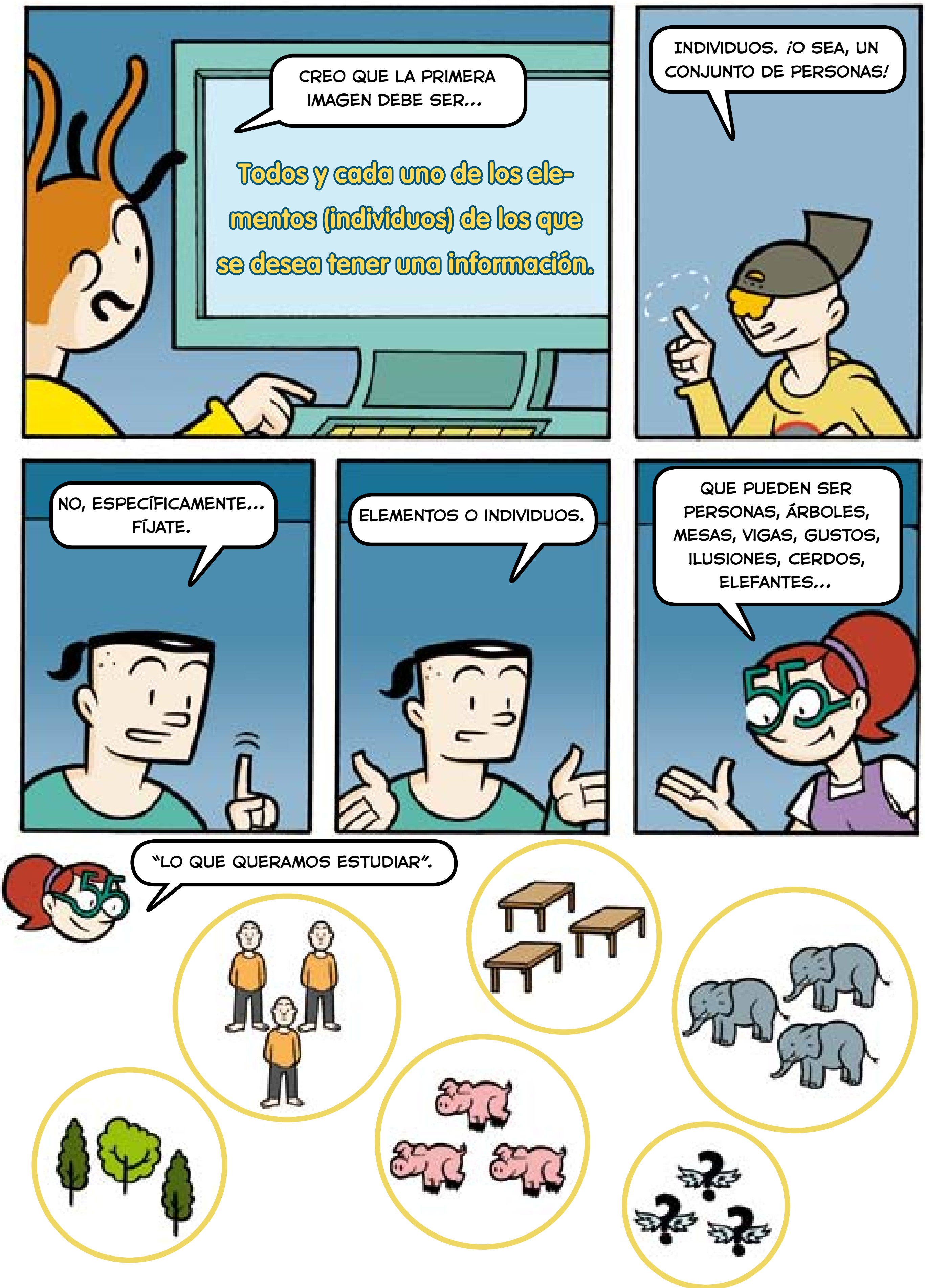
"La probabilidad de un elemento para cualquier distribución estadística de encontrarse entre la media y menos dos desviaciones típicas, y la media más dos desviaciones típicas es mayor al 75%".


PIENSO YOOO....
¡JA!
¡JA!
¡JA!
BUENO... QUIERO PENSAR QUE LA ESTADÍSTICA ME SERVIRÍA, POR EJEMPLO,
SI PONGO UNA FÁBRICA DE CAMISAS Y PANTALONES MODERNOS.
EXPLÍCATE... Y TE IREMOS PONIENDO LAS PEGAS.
¡NO TE DIGO!
BUENO, SI PUEDO HACER UNA ESTADÍSTICA DE ALTURAS Y DIMENSIONES DE TODOS LOS POSIBLES CLIENTES...
Y SI TAMBIÉN REALIZO UNA ESTADÍSTICA DE GUSTOS EN EL VESTIR...

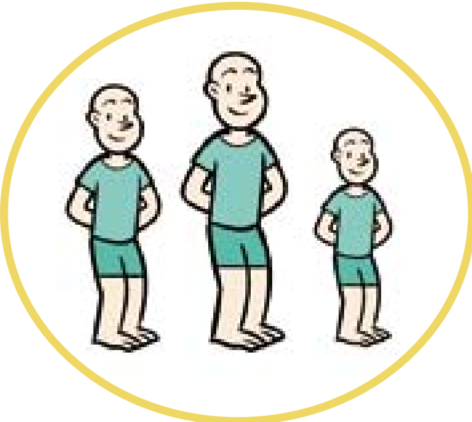

O SEA:
nada
ni sí ni no
un poco
sí
mucho
con locura
PODRÍA SABER QUÉ TALLAS FABRICAR Y EN QUÉ CANTIDAD CADA UNA DE ELLAS.
BUENO, HAS DICHO TANTAS COSAS, BASTANTE BUENAS, QUE CREO QUE DEBEMOS DEFINIRLAS CONCEPTUALMENTE, POR SU IMPORTANCIA.
VAMOS A LA BIBLIOTECA VIRTUAL, NOS INFORMAMOS, PREPARAMOS EL TEMA... A VER CÓMO SALE....
55


ID DÁNDOME MATERIAL E IRÉ FABRICANDO UNA PRESENTACIÓN.
DEBEREMOS REUNIR TODAS LAS IMÁGENES, VER LAS REPETICIONES...
ELEGIR Y DETERMINAR LAS QUE FALTEN.
PERO, PRIMERO HAY QUE REALIZARLAS...
YO HARÉ UN FILTRO DE LAS IMÁGENES ELABORADAS.



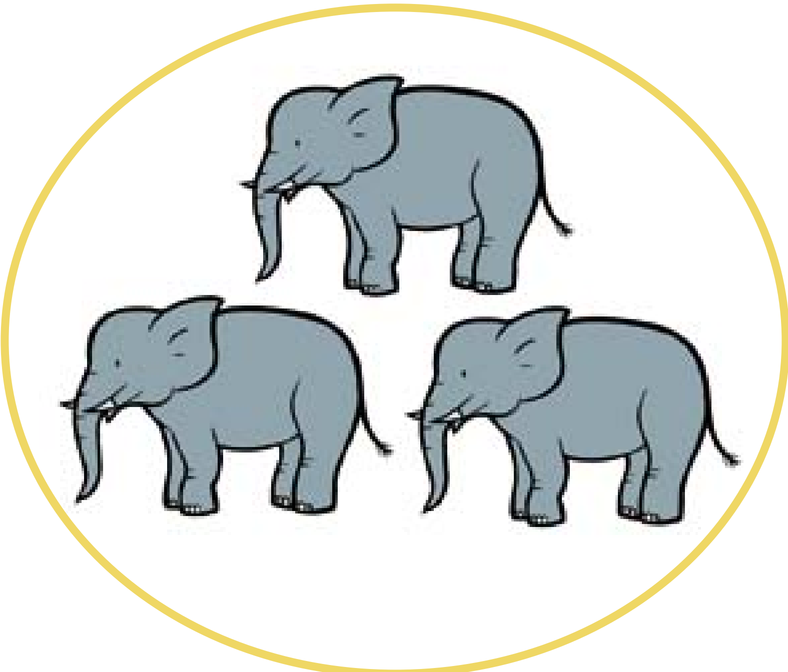

CREO QUE LA PRIMERA IMAGEN DEBE SER...
Todos y cada uno de los elementos (individuos) de los que se desea tener una información.
INDIVIDUOS. ¡O SEA, UN CONJUNTO DE PERSONAS!
NO, ESPECÍFICAMENTE... FÍJATE.
ELEMENTOS O INDIVIDUOS.
QUE PUEDEN SER PERSONAS, ÁRBOLES, MESAS, VIGAS, GUSTOS, ILUSIONES, CERDOS, ELEFANTES...
"LO QUE QUERAMOS ESTUDIAR".

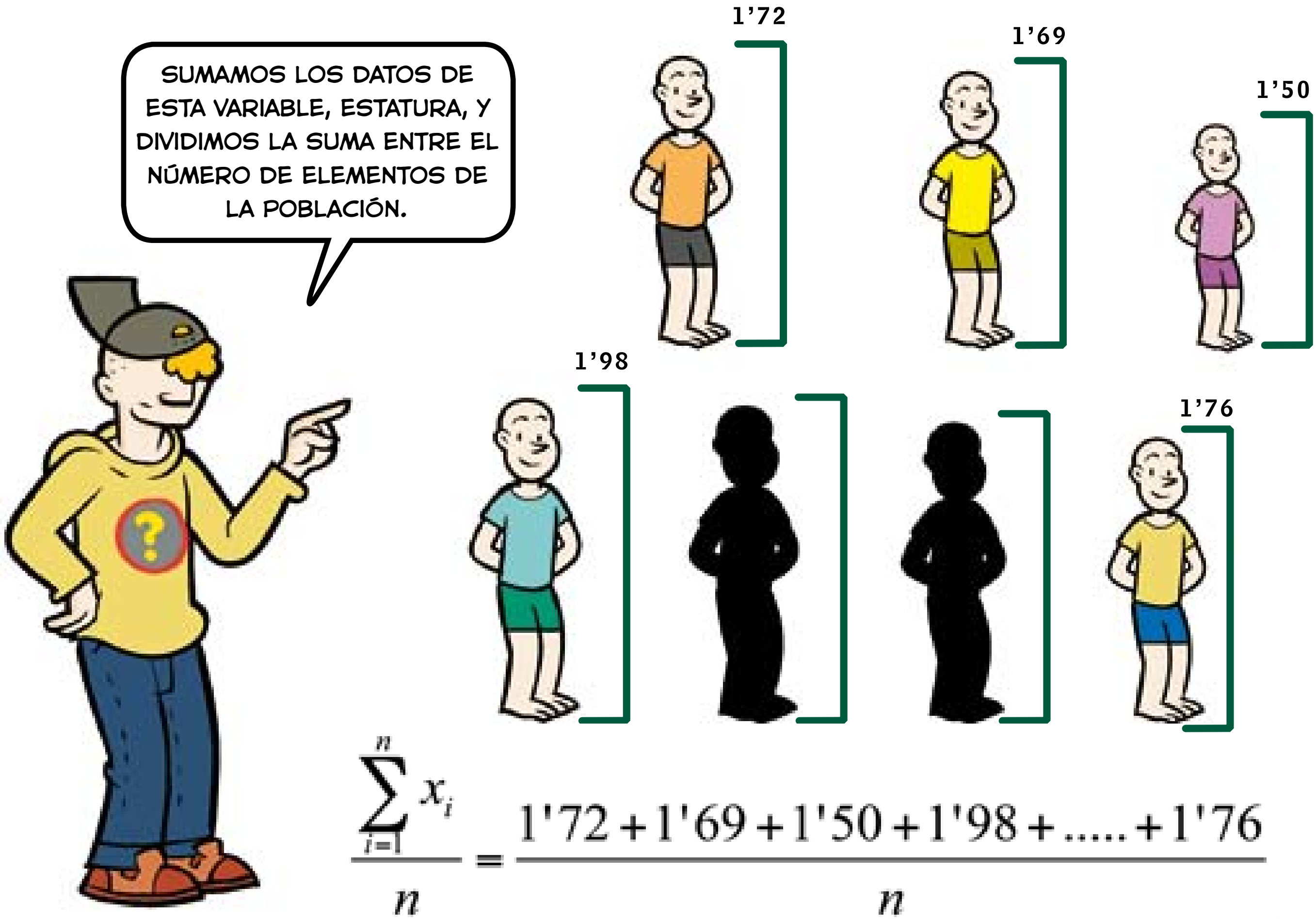
Estatura (cuantitativa)



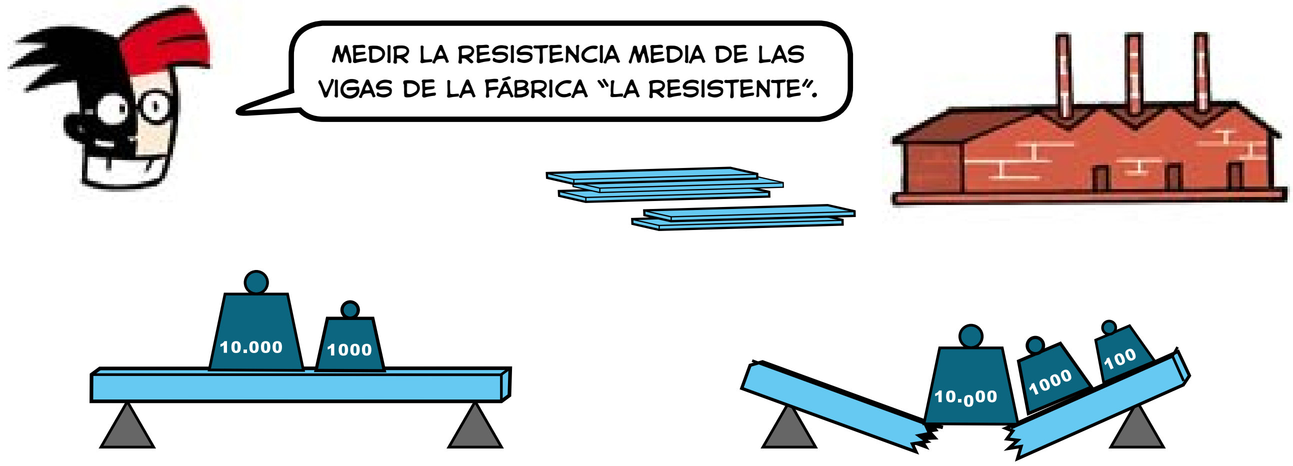
Elegancia (cualitativa)

Peso (cuantitativa)



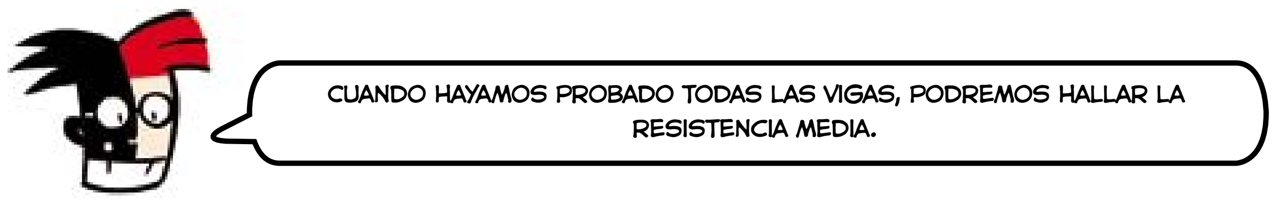


$$\frac{\sum_{i=1}^{n} x_i}{n} = \frac{1'72 + 1'69 + 1'50 + 1'98 + \ldots\ldots + 1'76}{n}$$

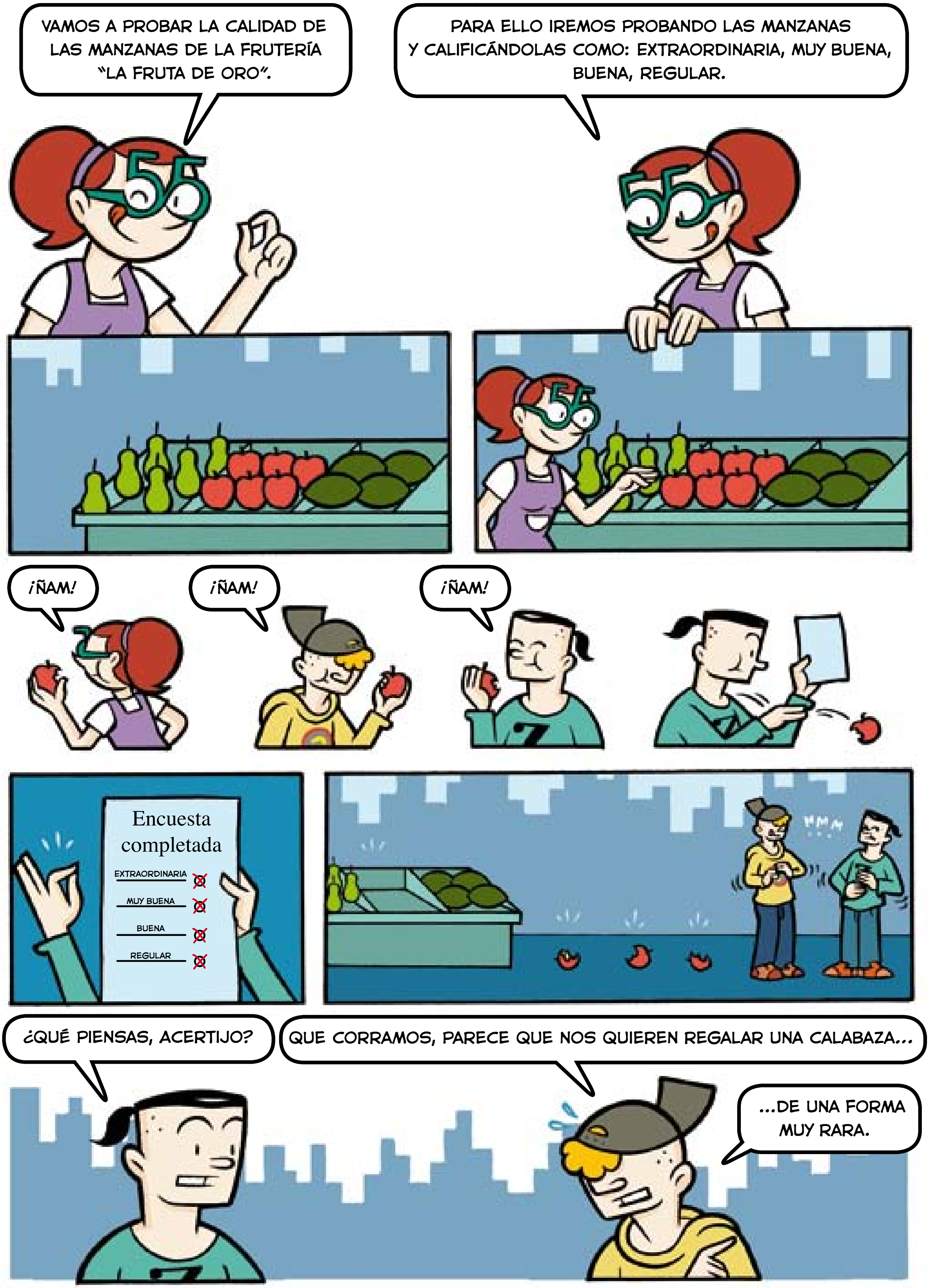


A VER, PONEMOS PESAS HASTA QUE SE ROMPA LA VIGA, Y ASÍ SABREMOS LO QUE RESISTEN.

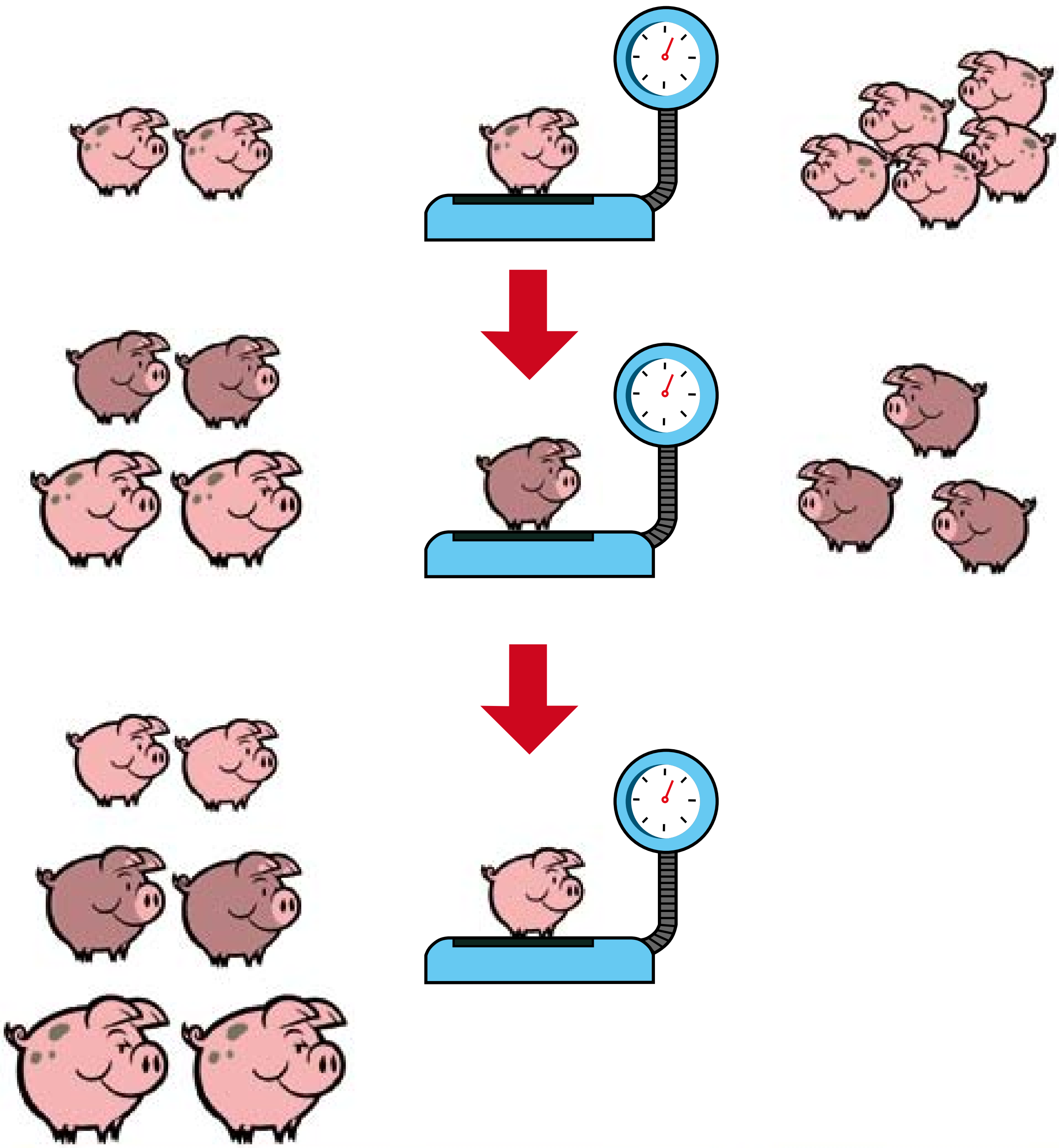
EN EL MOMENTO EN QUE SE ROMPE, APUNTAMOS EL PESO COLOCADO Y ESA SERÁ LA RESISTENCIA.



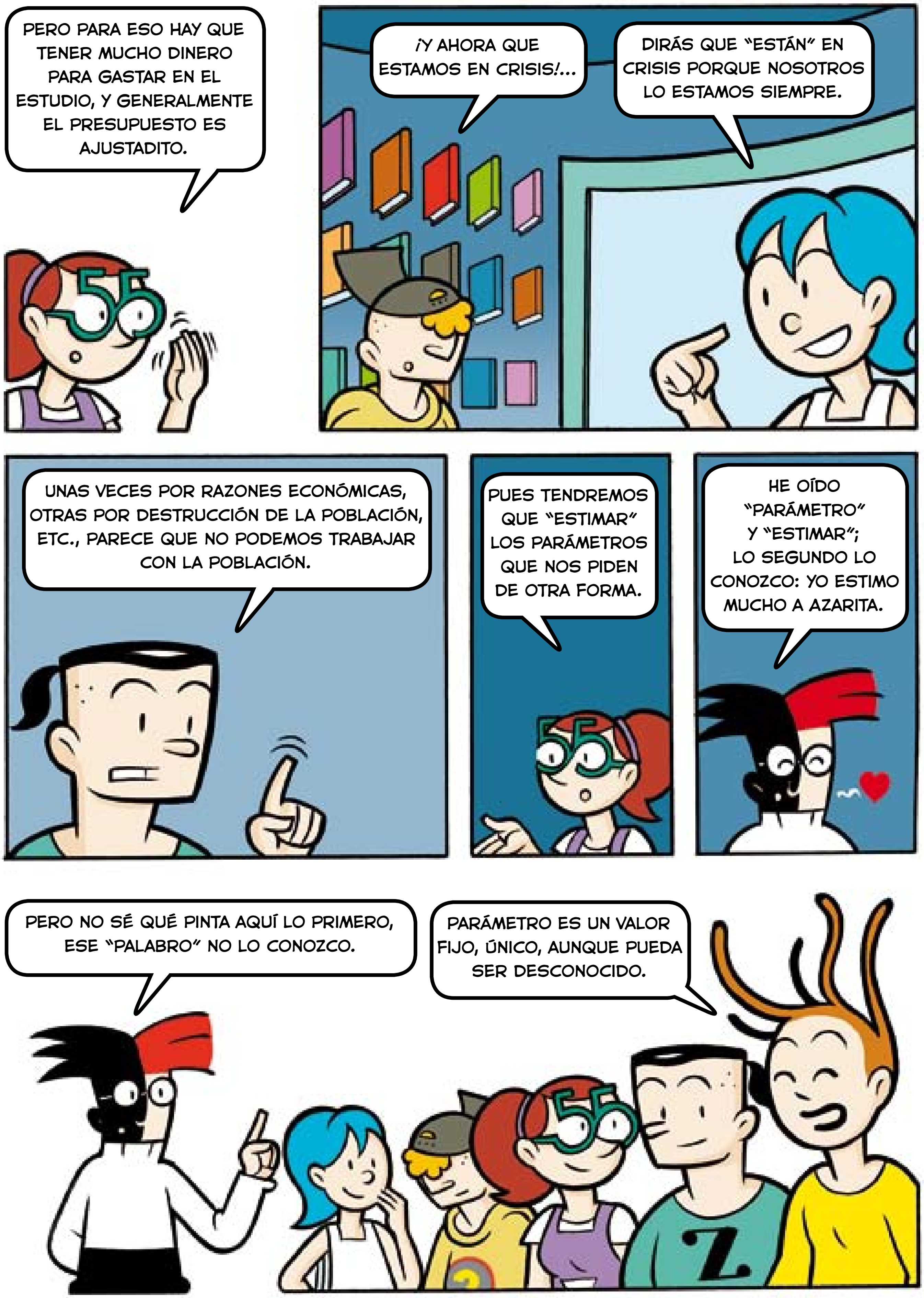
$$\frac{\sum_{i=1}^{n} x_i}{n} = \frac{532\,kg + 612\,kg + 586\,kg + 601\,kg + \ldots\ldots + 599\,kg}{100.000}$$

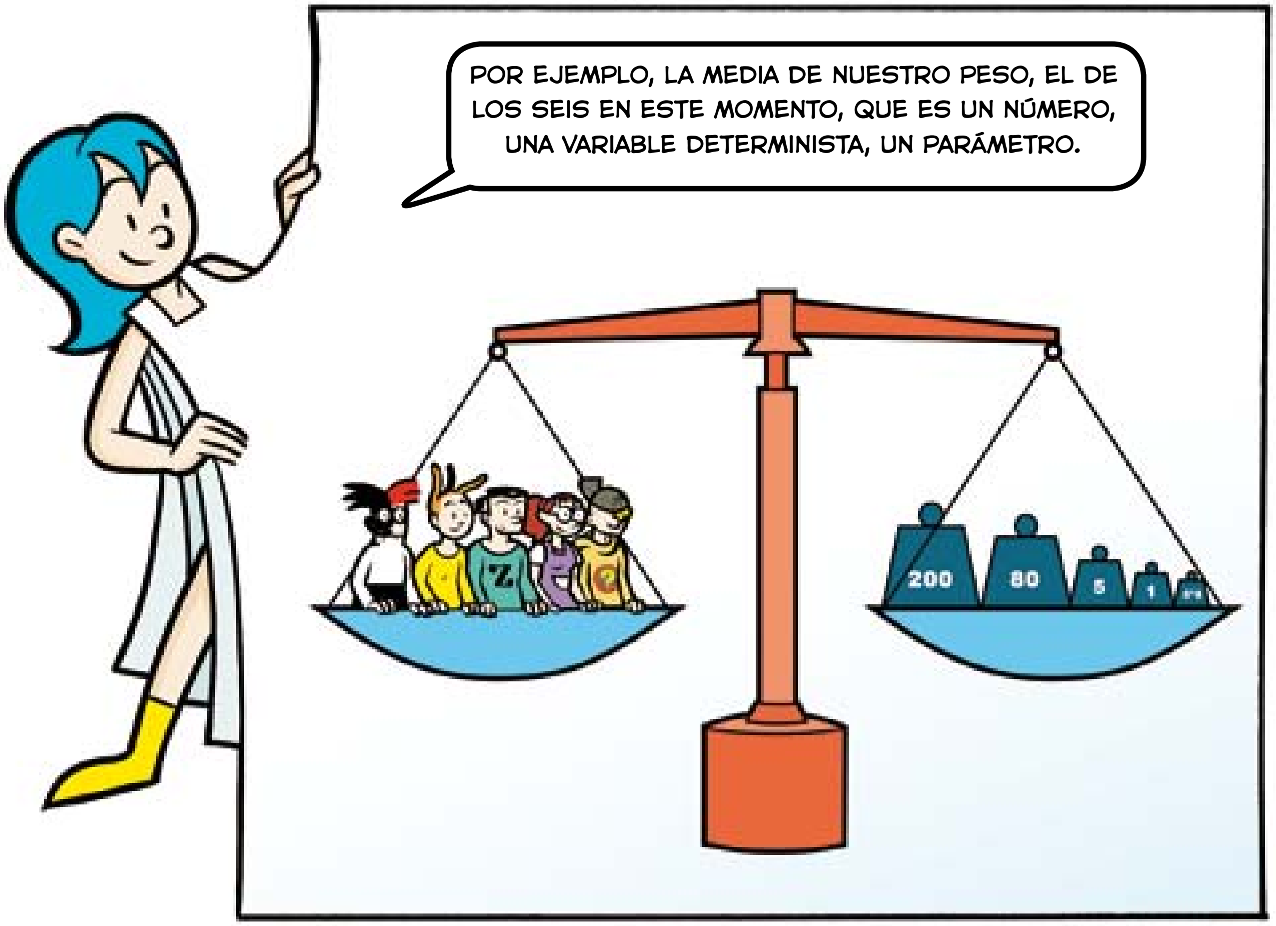

VAMOS A PROBAR LA CALIDAD DE LAS MANZANAS DE LA FRUTERÍA "LA FRUTA DE ORO".
PARA ELLO IREMOS PROBANDO LAS MANZANAS Y CALIFICÁNDOLAS COMO: EXTRAORDINARIA, MUY BUENA, BUENA, REGULAR.
¡ÑAM!
¡ÑAM!
¡ÑAM!
Encuesta completada
EXTRAORDINARIA
MUY BUENA
BUENA
REGULAR
¿QUÉ PIENSAS, ACERTIJO?
QUE CORRAMOS, PARECE QUE NOS QUIEREN REGALAR UNA CALABAZA...
...DE UNA FORMA MUY RARA.


BUENO, VAMOS A VER SI EL MÍO NO ACABA MAL. EJEMPLO NÚMERO 3:
ME HAN PEDIDO QUE HALLE EL PESO MEDIO DE TODOS LOS CERDOS DE LA EMPRESA "PORQUERIZAS EL GORRINO".
YA SE CUÁL ES LA POBLACIÓN... PERO LOS CERDOS, ¿DE CUÁNTAS PATAS SON?
DE CUATRO, HOMBRE... DE CUATRO. LOS IRÉ PESANDO UNO A UNO, Y CALCULANDO POSTERIORMENTE EL PESO MEDIO.




¡UF! VOY A TENER QUE REPETIR TODO EL PROCESO, PUES PARECE QUE LOS PRIMEROS QUE PESÉ HAN ENGORDADO.
¿QUE HAN ENGORDADO?... SI PARECEN ELEFANTES.
QUIZÁS SE PODRÍA ARREGLAR; TENIENDO MUCHÍSIMAS BALANZAS Y MUCHÍSIMOS PESADORES.


PERO PARA ESO HAY QUE TENER MUCHO DINERO PARA GASTAR EN EL ESTUDIO, Y GENERALMENTE EL PRESUPUESTO ES AJUSTADITO.
¡Y AHORA QUE ESTAMOS EN CRISIS!...
DIRÁS QUE "ESTÁN" EN CRISIS PORQUE NOSOTROS LO ESTAMOS SIEMPRE.
UNAS VECES POR RAZONES ECONÓMICAS, OTRAS POR DESTRUCCIÓN DE LA POBLACIÓN, ETC., PARECE QUE NO PODEMOS TRABAJAR CON LA POBLACIÓN.
PUES TENDREMOS QUE "ESTIMAR" LOS PARÁMETROS QUE NOS PIDEN DE OTRA FORMA.
HE OÍDO "PARÁMETRO" Y "ESTIMAR"; LO SEGUNDO LO CONOZCO: YO ESTIMO MUCHO A AZARITA.
PERO NO SÉ QUÉ PINTA AQUÍ LO PRIMERO, ESE "PALABRO" NO LO CONOZCO.
PARÁMETRO ES UN VALOR FIJO, ÚNICO, AUNQUE PUEDA SER DESCONOCIDO.

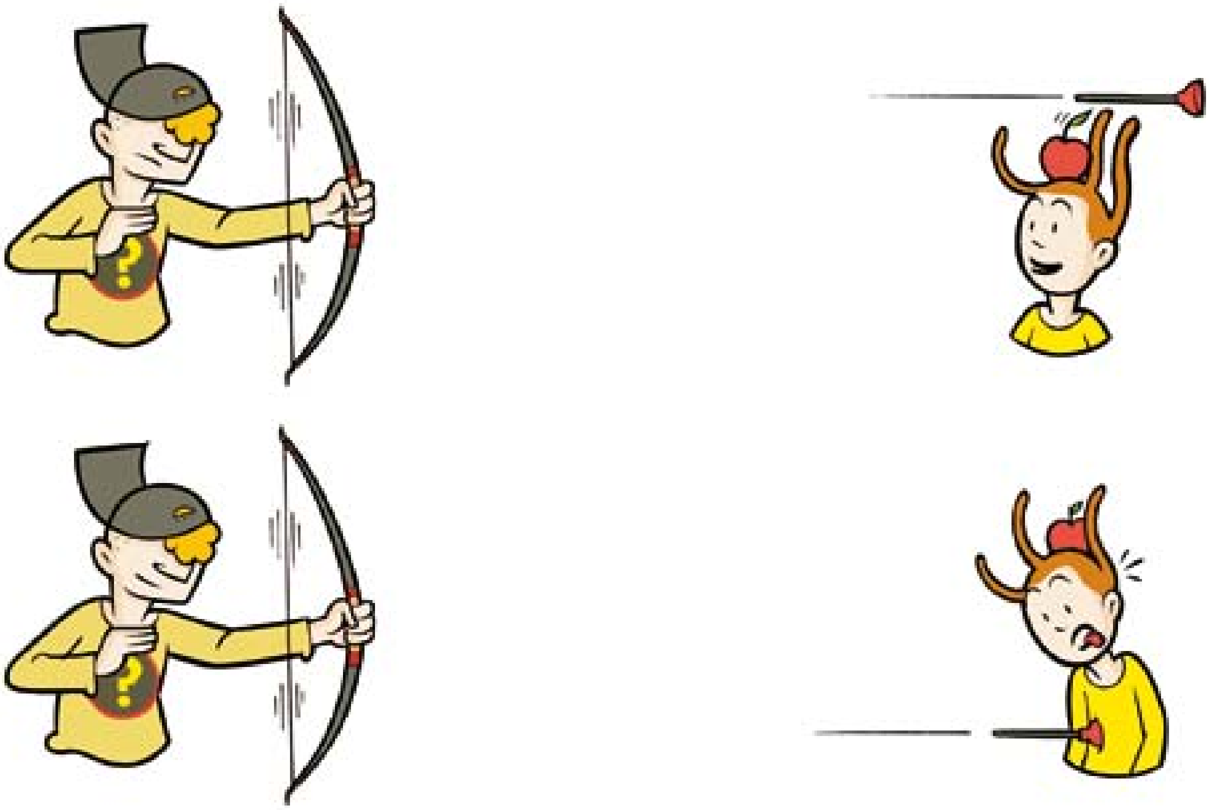

POR EJEMPLO, LA MEDIA DE NUESTRO PESO, EL DE LOS SEIS EN ESTE MOMENTO, QUE ES UN NÚMERO, UNA VARIABLE DETERMINISTA, UN PARÁMETRO.
200
80
5
1


ES DECIR, SUMAMOS NUESTROS PESOS Y LO DIVIDIMOS ENTRE 6; EL RESULTADO ES 47,8 KILOS, QUE ES EL PARÁMETRO PESO MEDIO DE LA POBLACIÓN FORMADA POR NOSOTROS SEIS.
Y DE LO DE "ESTIMAR", ¿QUÉ?... A MÍ ME INTERESA MUCHO.


UNA ESTIMACIÓN, YA LO DIJIMOS, ES UNA APROXIMACIÓN MÁS O MENOS ACERTADA DE UN PARÁMETRO, O UN INTERVALO DONDE PENSAMOS QUE SE ENCUENTRA EL PARÁMETRO.
¿QUÉ?
SI TÚ DICES: "DE AQUÍ HASTA ALLÍ, HABRÁ UNOS 15 METROS". NADIE PENSARÁ QUE SON 15 METROS JUSTOS (PARÁMETRO), SINO QUE APROXIMADAMENTE LOS HABRÁ, O SEA, UNA ESTIMACIÓN.
O SI DICES TENGO ENTRE 20 Y 25 EUROS, ES UNA ESTIMACIÓN, ES UNA VARIABLE ALEATORIA.
TE MIRAS LOS BOLSILLOS, CUENTAS, TIENES 21 €, ESTO ES UN PARÁMETRO, UNA VARIABLE DETERMINISTA.
¡AHORA LO PILLO!


LLEGAMOS A LA CONCLUSIÓN DE QUE O BIEN NO SE PUEDE CALCULAR EL PARÁMETRO MEDIANTE LA POBLACIÓN, O SE PUEDE CON MUCHAS DIFICULTADES, GENERALMENTE.
POR LO QUE TENDREMOS QUE INVENTAR ALGO QUE NOS DÉ UNA APROXIMACIÓN, A SER POSIBLE MUY BUENA, SOBRE EL PARÁMETRO, "UN ESTIMADOR".
EXISTEN DOS CLASES DE ESTIMADORES: PUNTUALES Y POR INTERVALOS.
ANTES DE SEGUIR OS QUIERO RECORDAR UNA ESCENA DE NUESTROS ANTERIORES EPISODIOS SOBRE ESTIMACIÓN.
ESTIMO QUE ACERTIJO MIDE 1,80 MS. O MEJOR ENTRE 1,75 Y 2,10 MS. DE ALTURA, Y QUE GRÁFICA SE ENCUENTRA ENTRE 2 Y 2,50 MS. DE ALTURA.
MIDO... 1,79 MS.
EL VALOR DE LA MEDIDA ES VARIABLE DETERMINISTA.
LA ESTIMACIÓN, YA SEA PUNTUAL O POR INTERVALO, ES UNA VARIABLE ALEATORIA O ESTOCÁSTICA.

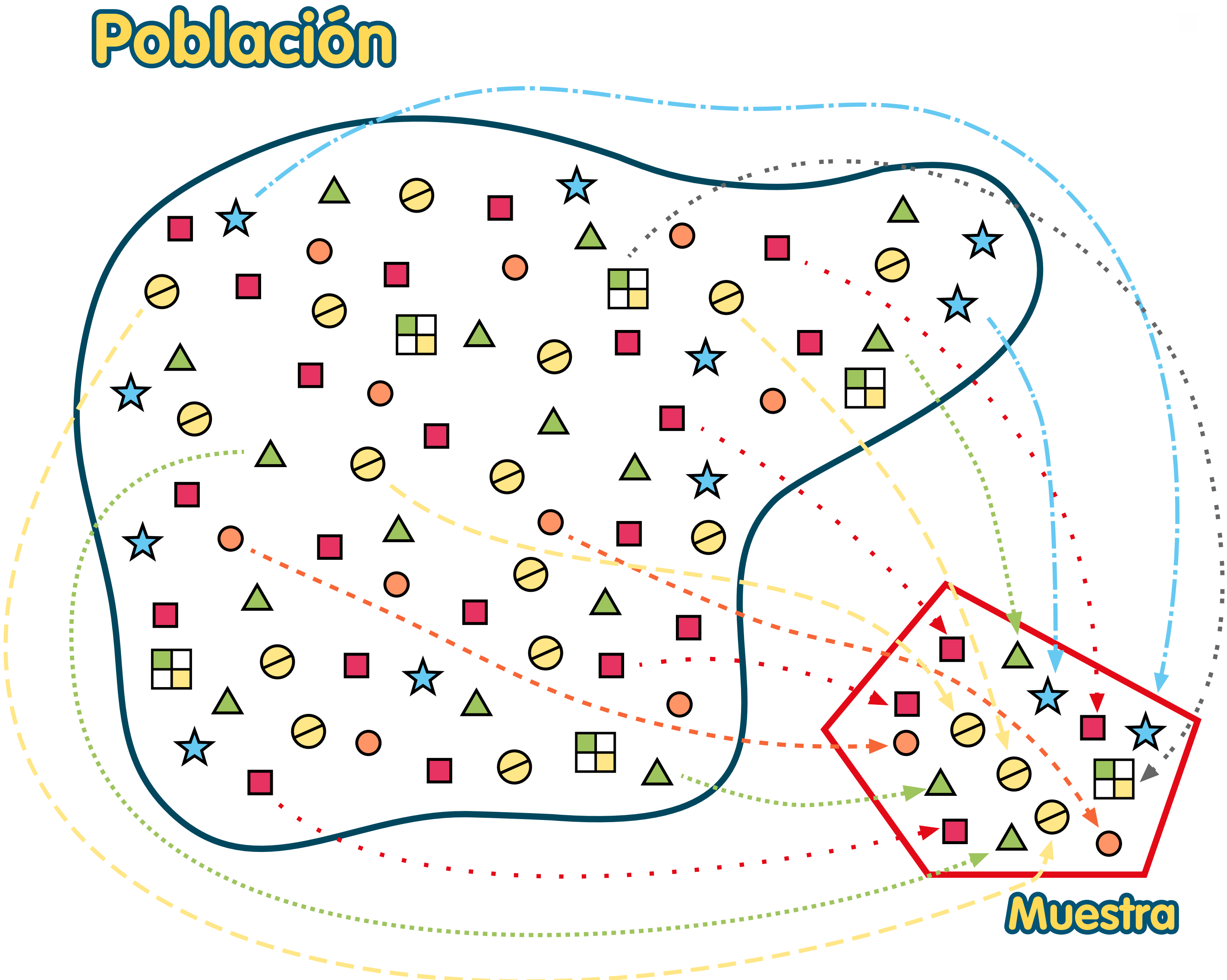
Más vale acertar aproximadamente que fallar exactamente.




LAS FRASES SON MUY BONITAS..., PERO ¿A QUÉ VIENEN?
LA PRIMERA ES PARA QUE NOS ACORDEMOS DE QUE CUANDO HACEMOS UNA ESTIMACIÓN SIEMPRE CORREMOS UN RIESGO DE EQUIVOCARNOS, DE COMETER UN ERROR.
SUPONGO QUE YA SE HABRÁN HECHO INVESTIGACIONES PARA QUE SEA LO MÁS PEQUEÑO POSIBLE.
UNA COSA PARECIDA, PODÍAMOS DECIR QUE ES VERDAD CON UN CIERTO RIESGO.
LA SEGUNDA NOS INDICA QUE LAS ESTIMACIONES DEBEMOS HACERLAS POR INTERVALOS, NUNCA PUNTUALMENTE, PUES CORREMOS UN ELEVADO RIESGO DE FALLAR.
PUES SIEMPRE DAREMOS EN TODA ESTIMACIÓN UN INTERVALO, BUENO...
CUANDO SEPAMOS HACERLO.


¡ALEGRÍA! ¡ALEGRÍA!
OS LO DIGO MÁS CLARO, SEGUID CREANDO PROBLEMAS.
¿PERO DE QUÉ PROBLEMAS HABLAS?
LOS ENUMERARÉ:
1º: ¿Cómo hallamos el estimador?
2º: ¿Dónde lo buscamos?
3º: ¿Cómo definimos el intervalo?
4º: ¿Cómo comprobamos que la estimación es aceptable?
5º: ¿Qué utilidad le daremos?
PARA ELLO VAMOS A ANALIZAR VARIAS CUESTIONES QUE AL FINAL CREO QUE TE DEJARÁN CONTENTO.
DE ACUERDO, TENDRÉ UN POCO DE PACIENCIA, PERO POR FAVOR, NO LO COMPLIQUEMOS DEMASIADO.

Muestra: Conjunto reducido de elementos de una población, extraídos convenientemente para un mantenimiento lo más proporcional posible, con respecto a aquella, de las variables o características estudiadas; de dicha muestra obtendremos una información, que podremos inferir respecto de la población.



DIMENSIÓN MUESTRAL= 15

ME GUSTARÍA EXPLICAROS CUATRO COSAS SOBRE EL ESQUEMA:

1º: La población sería mucho mayor, pero por claridad del dibujo vamos a suponer que su tamaño muestral es 75.

2º: Se ha decidido que la muestra sea de 15 elementos.

3º: La muestra se elige por unos procedimientos que supongo que veremos posteriormente; aquí la elección es teórica, para poder establecer las definiciones básicas.

YO HE ENCONTRADO QUE:

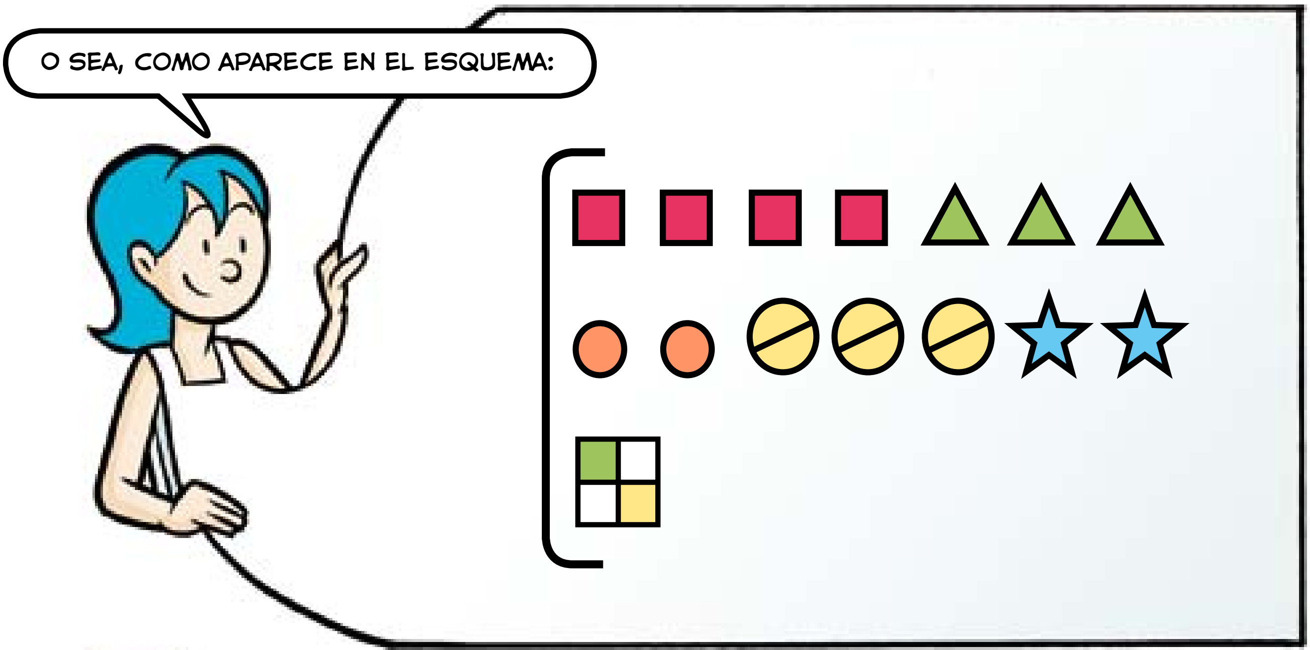
Unidad Muestral: Es cada uno de los posibles componentes de la muestra.

PUES YO HE ENCONTRADO QUE LO QUE HACÍAMOS ANTES SE LLAMA "CENSO" PORQUE:

Censo: El estudio de todos y cada uno de los elementos de la población.

YO APORTO LO SIGUIENTE:

Marco muestral: Enumeración exhaustiva de todas las unidades muestrales.

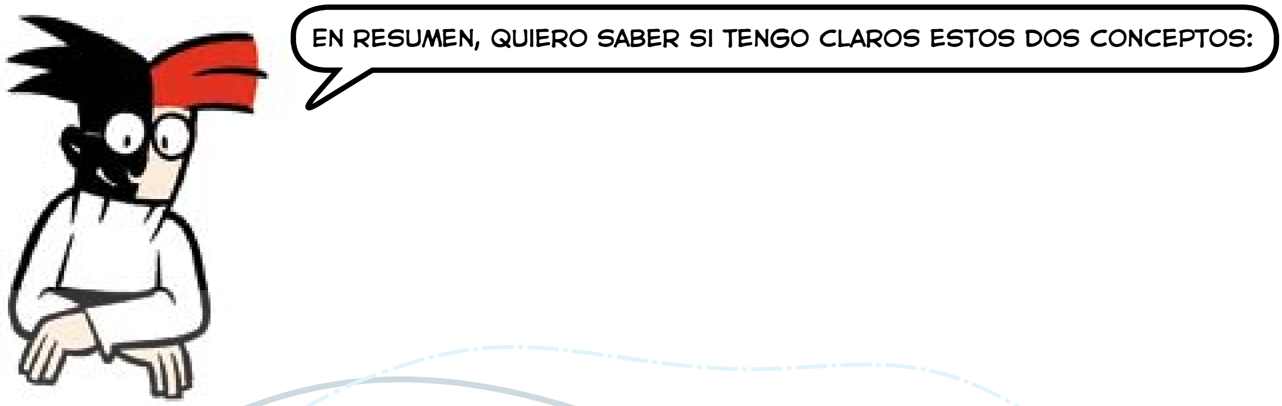


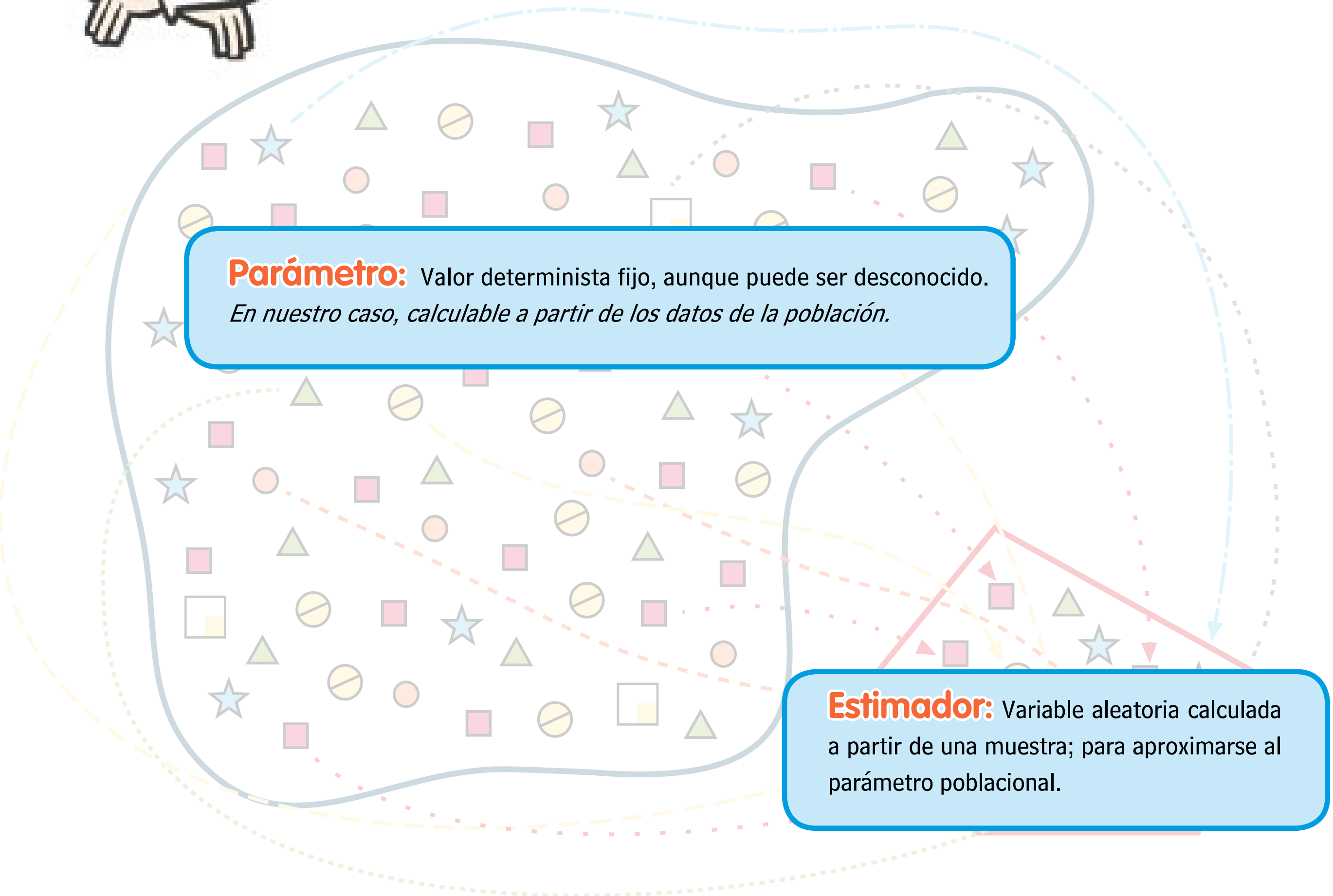
Población marco: Cada uno de los posibles componentes de la muestra.

Coeficiente de elevación: Cantidad de unidades poblacionales representadas por cada uno de los componentes de la muestra.

Fracción de muestreo: Inversa del coeficiente de elevación.

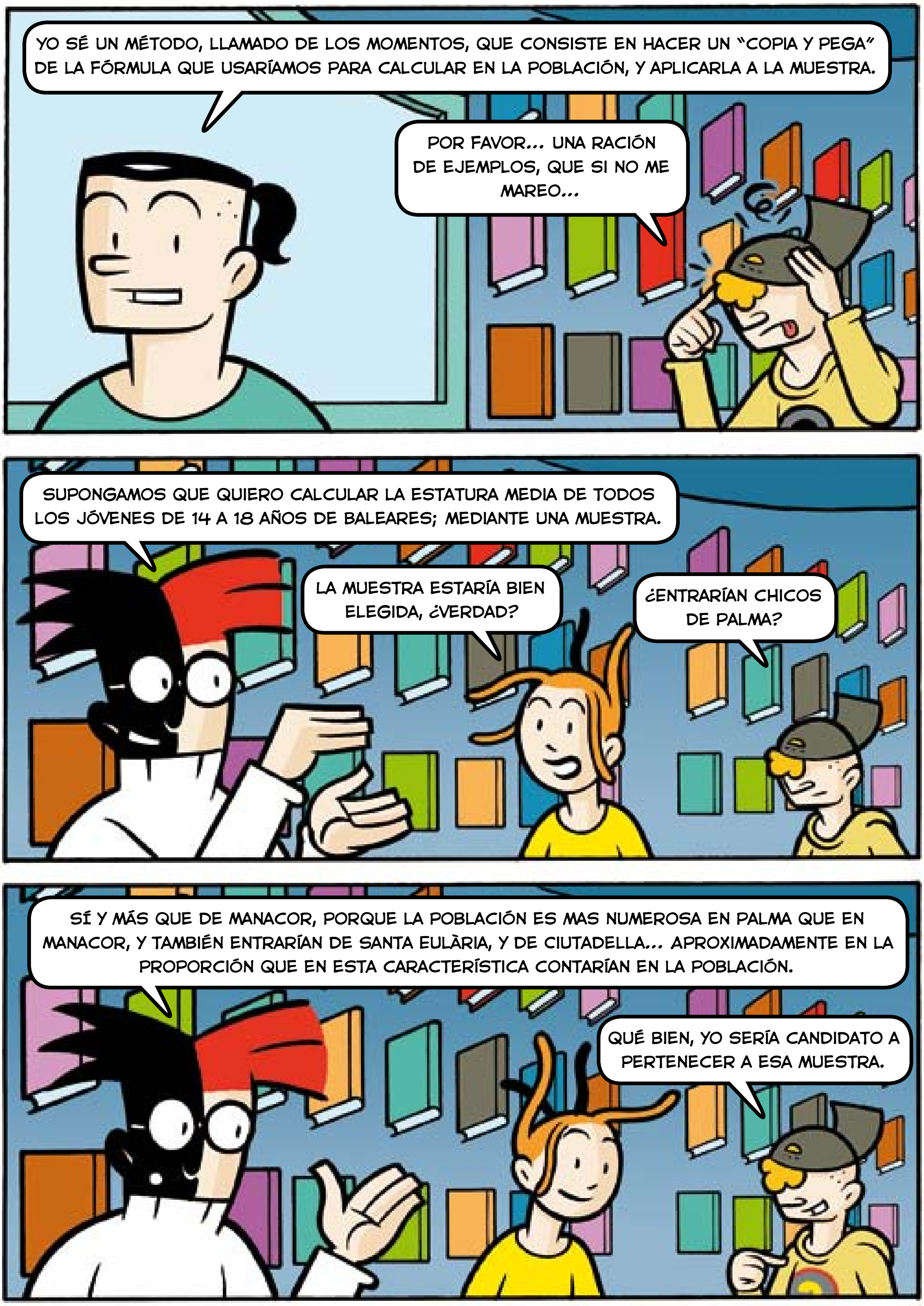
$$\frac{1}{5}$$





Parámetro: Valor determinista fijo, aunque puede ser desconocido.
En nuestro caso, calculable a partir de los datos de la población.

Estimador: Variable aleatoria calculada a partir de una muestra; para aproximarse al parámetro poblacional.


YO SÉ UN MÉTODO, LLAMADO DE LOS MOMENTOS, QUE CONSISTE EN HACER UN "COPIA Y PEGA" DE LA FÓRMULA QUE USARÍAMOS PARA CALCULAR EN LA POBLACIÓN, Y APLICARLA A LA MUESTRA.
POR FAVOR... UNA RACIÓN DE EJEMPLOS, QUE SI NO ME MAREO...
SUPONGAMOS QUE QUIERO CALCULAR LA ESTATURA MEDIA DE TODOS LOS JÓVENES DE 14 A 18 AÑOS DE BALEARES; MEDIANTE UNA MUESTRA.
LA MUESTRA ESTARÍA BIEN ELEGIDA, ¿VERDAD?
¿ENTRARÍAN CHICOS DE PALMA?
SÍ Y MÁS QUE DE MANACOR, PORQUE LA POBLACIÓN ES MAS NUMEROSA EN PALMA QUE EN MANACOR, Y TAMBIÉN ENTRARÍAN DE SANTA EULÀRIA, Y DE CIUTADELLA... APROXIMADAMENTE EN LA PROPORCIÓN QUE EN ESTA CARACTERÍSTICA CONTARÍAN EN LA POBLACIÓN.
QUÉ BIEN, YO SERÍA CANDIDATO A PERTENECER A ESA MUESTRA.

Población $\{x_1, x_2, x_3, \ldots, x_{36}, \ldots, x_{73}, x_{74}, x_{75}\}$

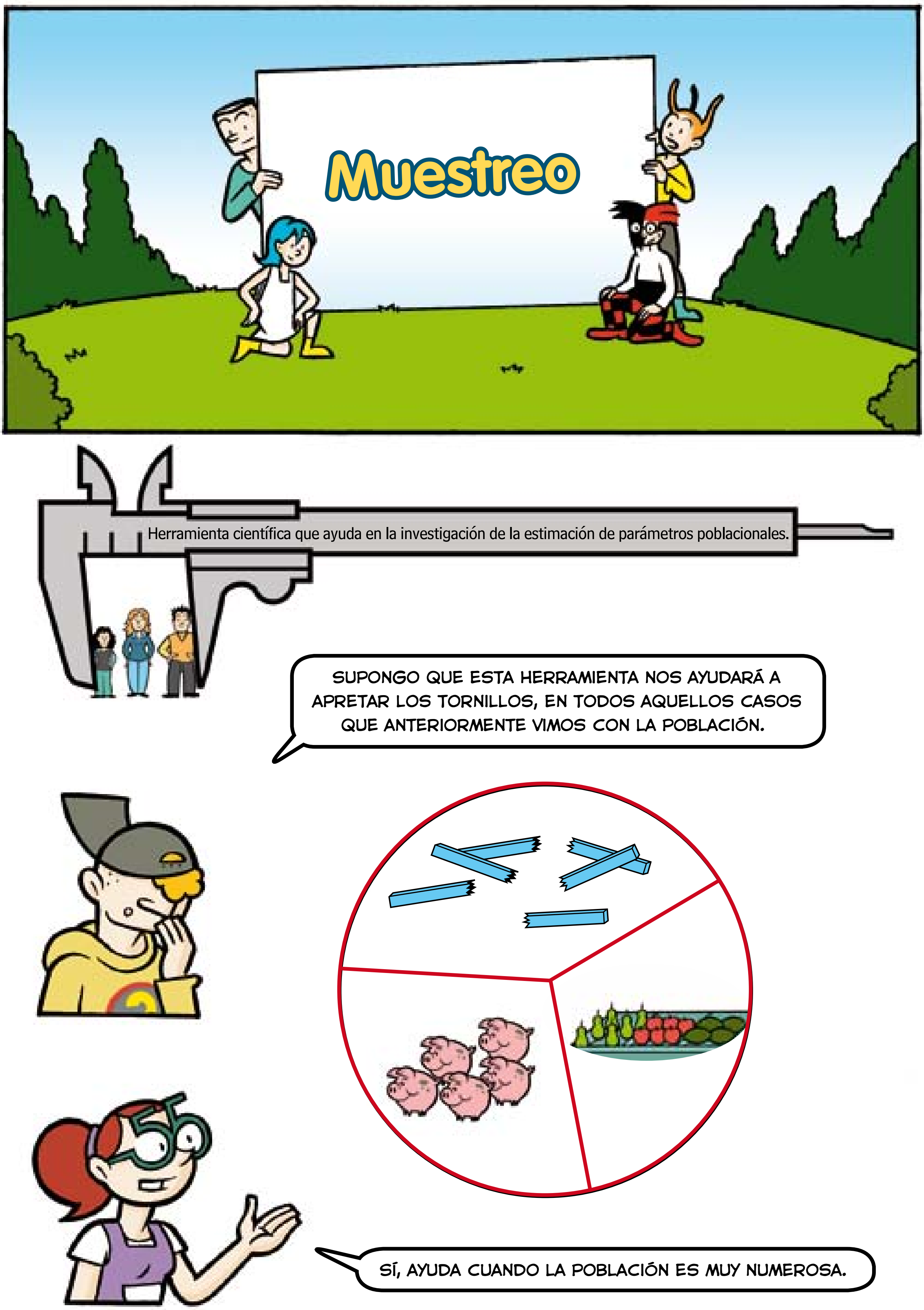
Media $= \mu = \dfrac{x_1 + x_2 + x_3 + \ldots\ldots + x_{36} + \ldots\ldots + x_{73} + x_{74} + x_{75}}{75}$

Muestra $\{x_3, x_7, x_{12}, x_{19}, x_{22}, x_{23}, x_{30}, x_{36}, x_{51}, x_{53}, x_{59}, x_{60}, x_{64}, x_{68}, x_{72}\}$

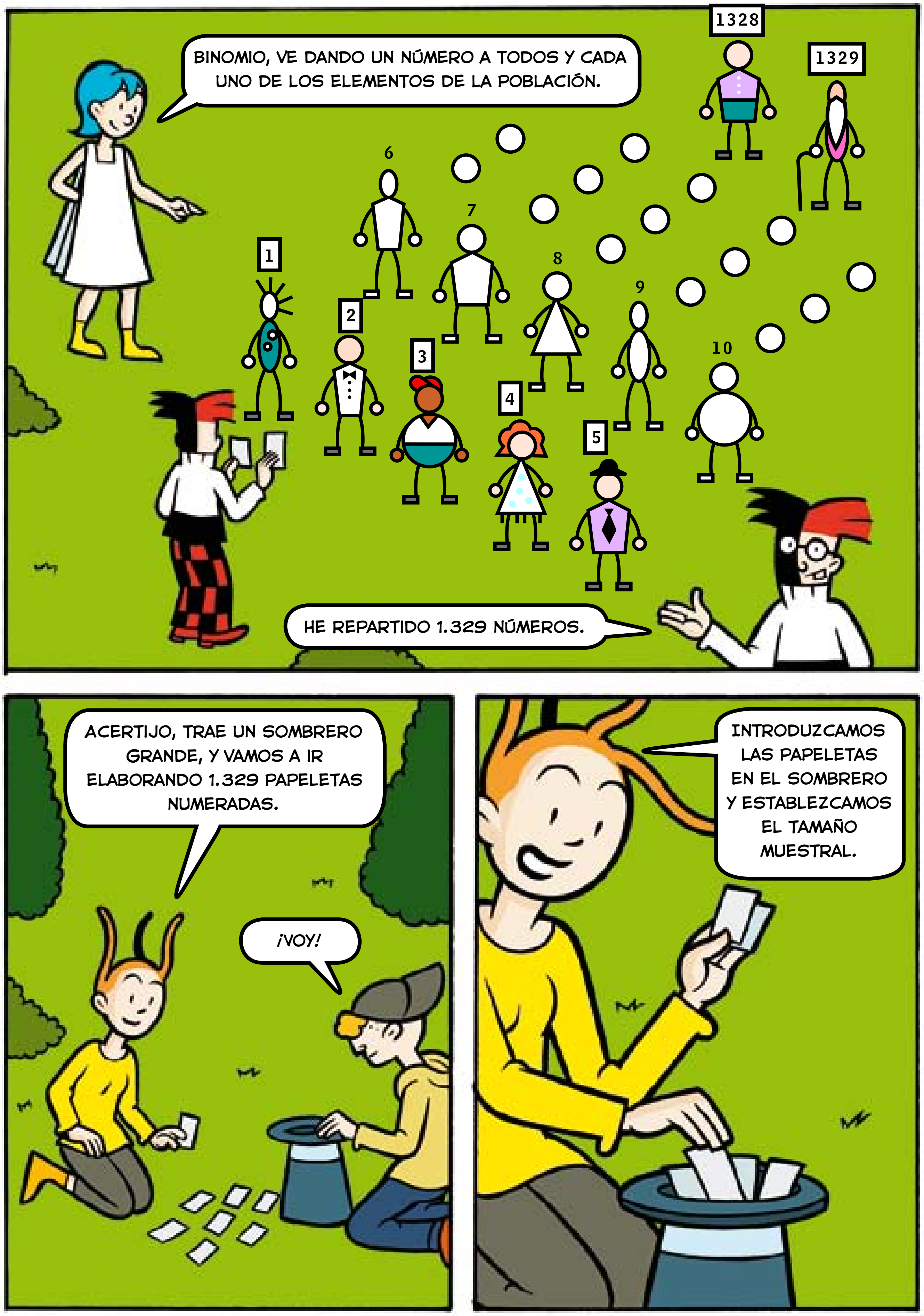
Estimación de la media $= \hat{\mu} = \overline{x}$

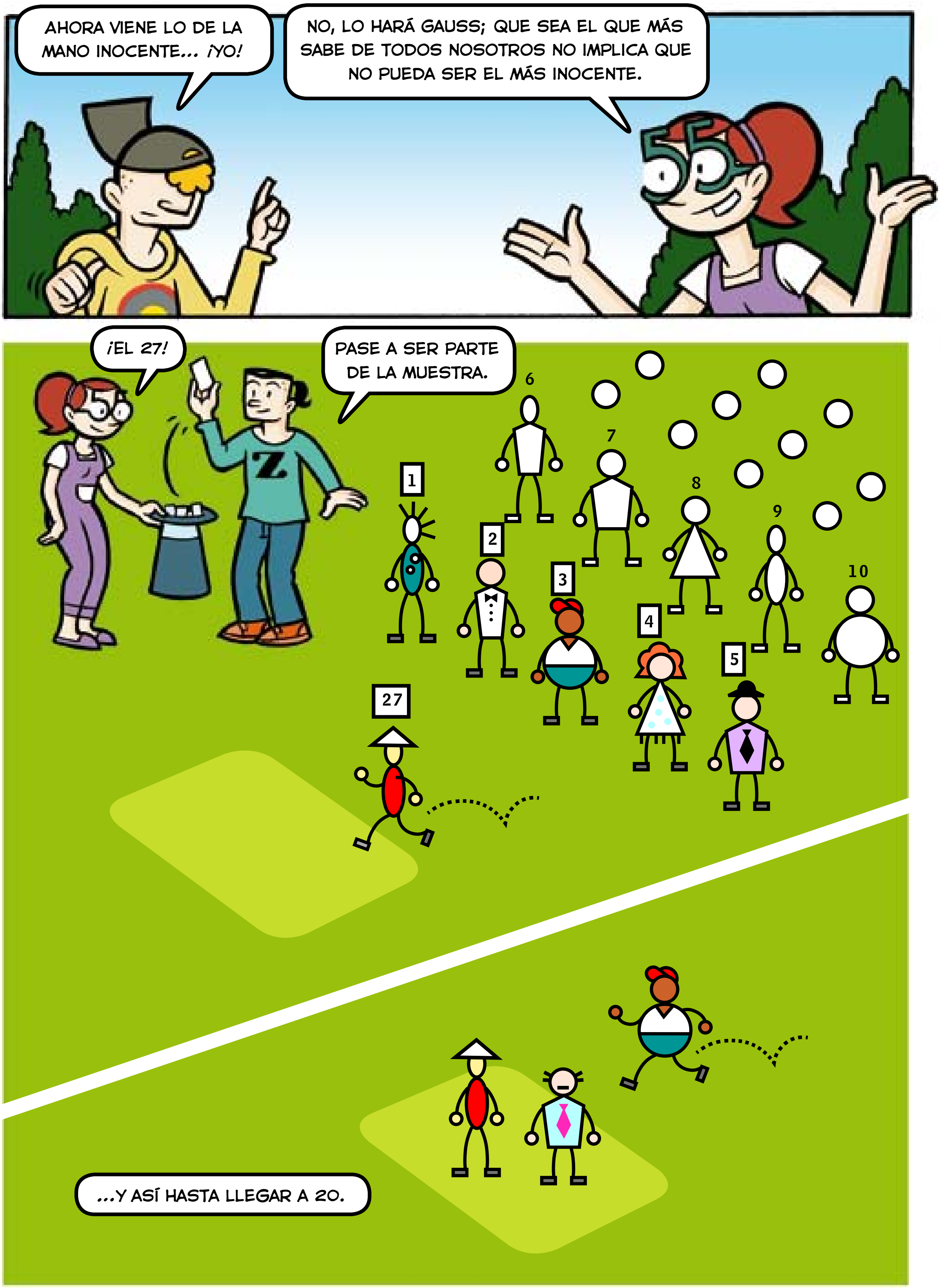
$$\hat{\mu} = \overline{x} = \frac{x_3 + x_7 + x_{12} + x_{19} + x_{22} + x_{23} + x_{30} + x_{36} + x_{51} + x_{53} + x_{59} + x_{60} + x_{64} + x_{68} + x_{72}}{15}$$


LOS INTERVALOS LOS VEREMOS HACIA EL FINAL, AHORA HAREMOS UNA PAUSA PARA ENFRENTARNOS A LOS DISTINTOS TIPOS DE MUESTREO.
PRIMERO HAREMOS UN LARGO DESCANSO, MUUUUUY LAAAAAARGO Y DESPUÉS VEREMOS LAS FORMAS DE MUESTREAR.
DE ACUERDOOOOO.
ZZZZZZ
¡¡ARRIBA!! ¡VAMOS CON EL MUESTREO!

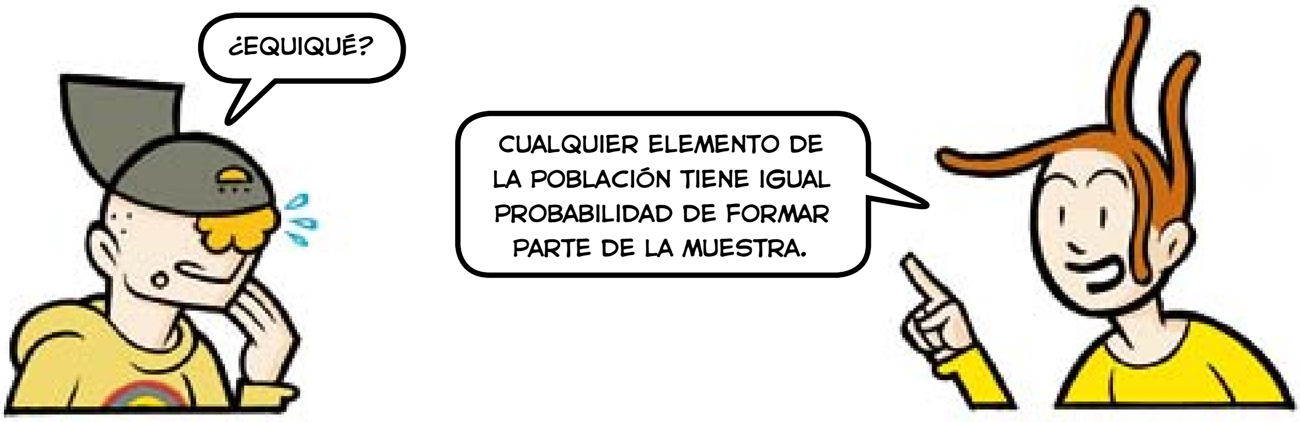

Muestreo
Herramienta científica que ayuda en la investigación de la estimación de parámetros poblacionales.
SUPONGO QUE ESTA HERRAMIENTA NOS AYUDARÁ A APRETAR LOS TORNILLOS, EN TODOS AQUELLOS CASOS QUE ANTERIORMENTE VIMOS CON LA POBLACIÓN.
SÍ, AYUDA CUANDO LA POBLACIÓN ES MUY NUMEROSA.


TAMBIÉN CUANDO LA MEDICIÓN HA DE HACERSE CON UN PROCEDIMIENTO DESTRUCTIVO.
CUANDO HAY PROCESOS DE ALTERACIÓN.
VEAMOS, PUES, LOS PRINCIPALES TIPOS DE MUESTREO:
PROBABILÍSTICOS
NO PROBABILÍSTICOS
PROBABILÍSTICOS
MAS:
Muestreo aleatorio simple.
Muestreo aleatorio sistemático.
Muestreo estratificado.
Muestreo por conglomerados.
Muestreo polietápico.
Muestreos especiales.
NO PROBABILÍSTICOS
Muestreo de conveniencia.
Muestreo de juicios.
Muestreo por cuotas.
Muestreo de bola de nieve.
EMPECEMOS VIENDO QUÉ ES ESE "MAS"... QUE DEBE SER LO MÁS.
ES EL MÉTODO MÁS SENCILLO CONOCIDO, PERO GENERALMENTE EL MENOS USADO.


BINOMIO, VE DANDO UN NÚMERO A TODOS Y CADA UNO DE LOS ELEMENTOS DE LA POBLACIÓN.
1328
1329
6
7
1
8
9
2
3
10
4
5
HE REPARTIDO 1.329 NÚMEROS.
ACERTIJO, TRAE UN SOMBRERO GRANDE, Y VAMOS A IR ELABORANDO 1.329 PAPELETAS NUMERADAS.
¡VOY!
INTRODUZCAMOS LAS PAPELETAS EN EL SOMBRERO Y ESTABLEZCAMOS EL TAMAÑO MUESTRAL.


AHORA VIENE LO DE LA MANO INOCENTE... ¡YO!
NO, LO HARÁ GAUSS; QUE SEA EL QUE MÁS SABE DE TODOS NOSOTROS NO IMPLICA QUE NO PUEDA SER EL MÁS INOCENTE.
¡EL 27!
PASE A SER PARTE DE LA MUESTRA.
1
2
3
4
5
6
7
8
9
10
27
...Y ASÍ HASTA LLEGAR A 20.


HAGAMOS ALGUNAS TRANSPARENCIAS DE DEFINICIÓN Y RESUMEN DEL PROCESO.
EL MAS O MUESTREO ALEATORIO SIMPLE, SELECCIONA LA MUESTRA POR PROCESOS ALEATORIOS, CUMPLIENDO LA CONDICIÓN DE SER EQUIPROBABILÍSTICO PARA TODO ELEMENTO.


¿EQUIQUÉ?
CUALQUIER ELEMENTO DE LA POBLACIÓN TIENE IGUAL PROBABILIDAD DE FORMAR PARTE DE LA MUESTRA.

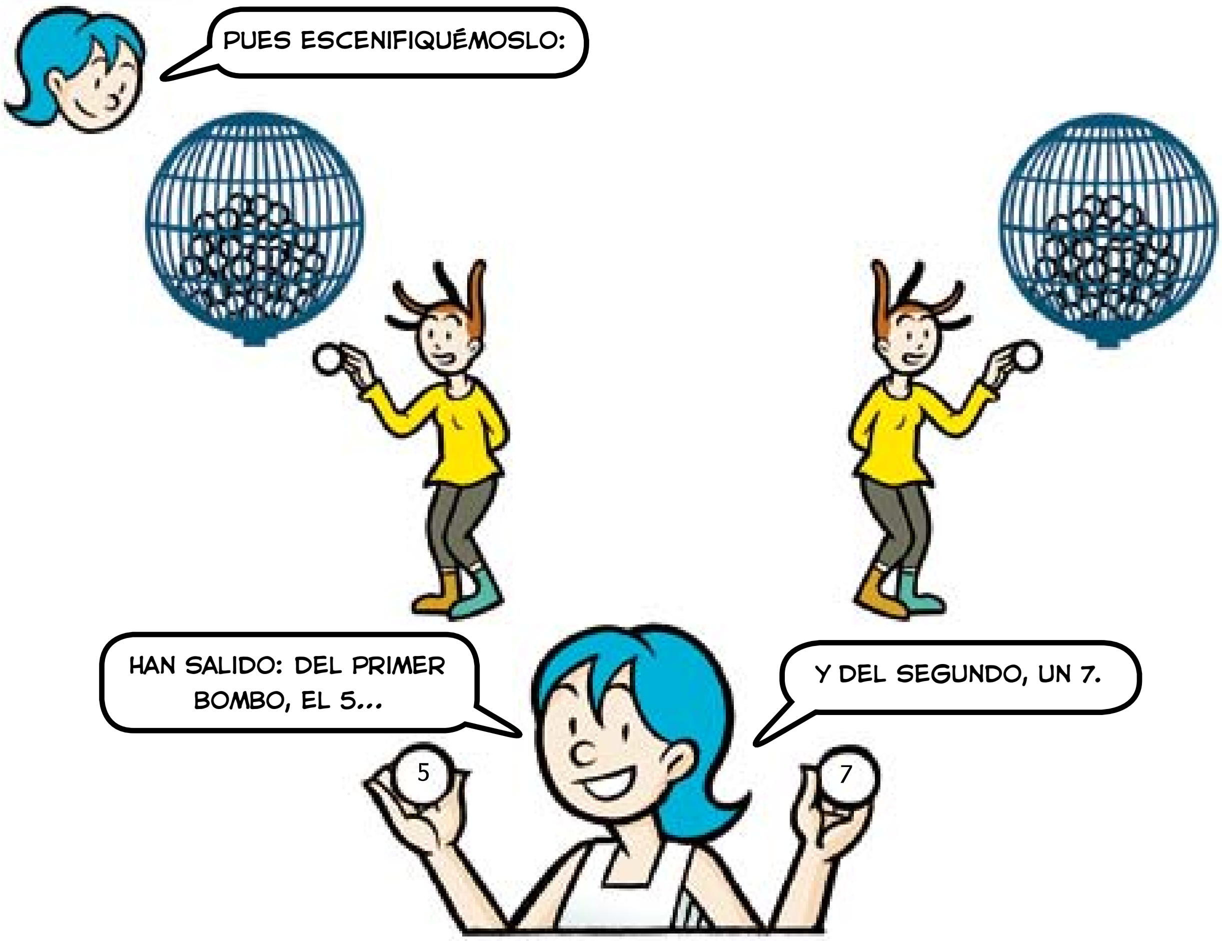

SIN REPOSICIÓN
CON REPOSICIÓN

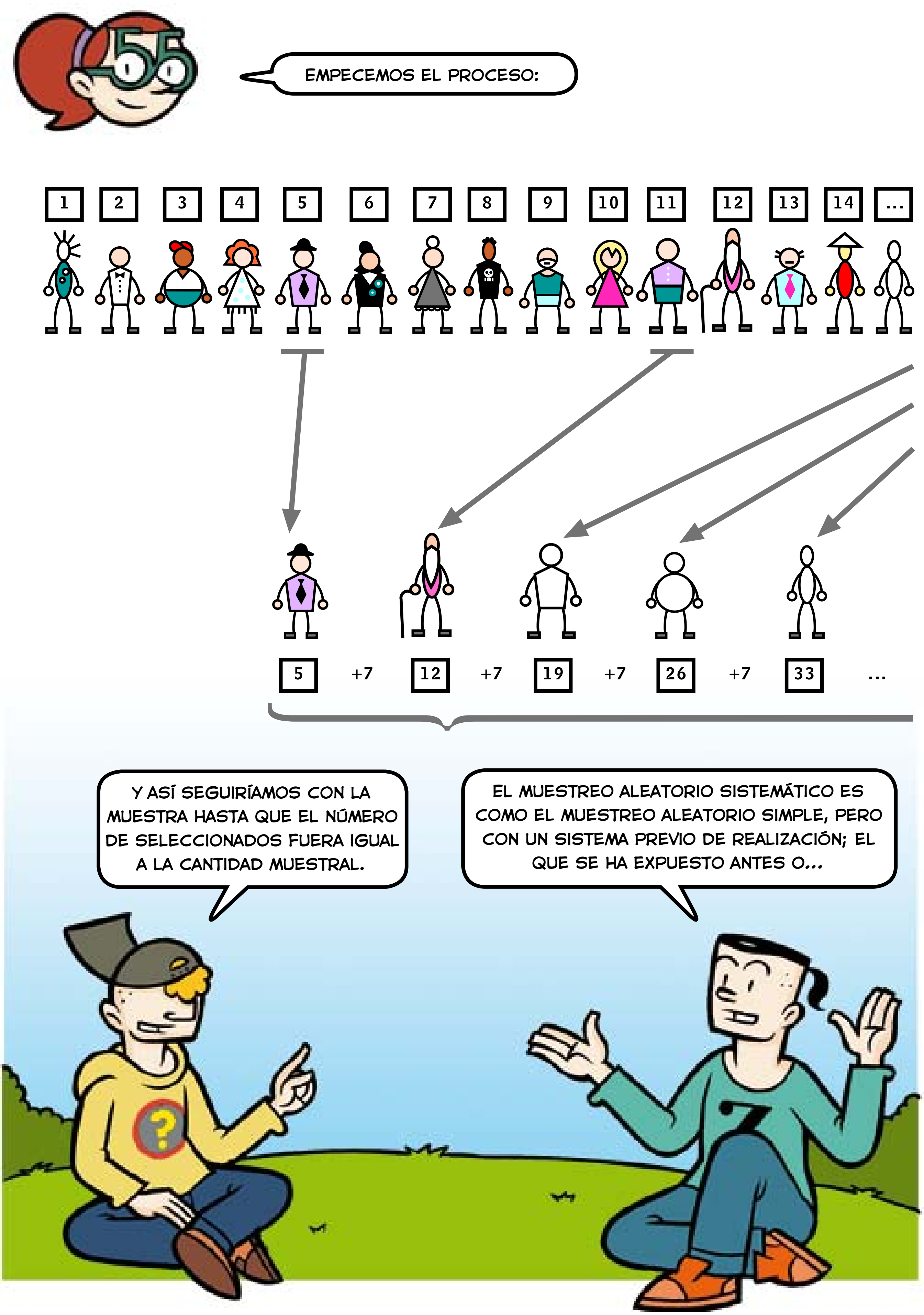

EL PROCEDIMIENTO PUEDE SER CON REPOSICIÓN Y SIN REPOSICIÓN.
¿CUÁL ES MEJOR?

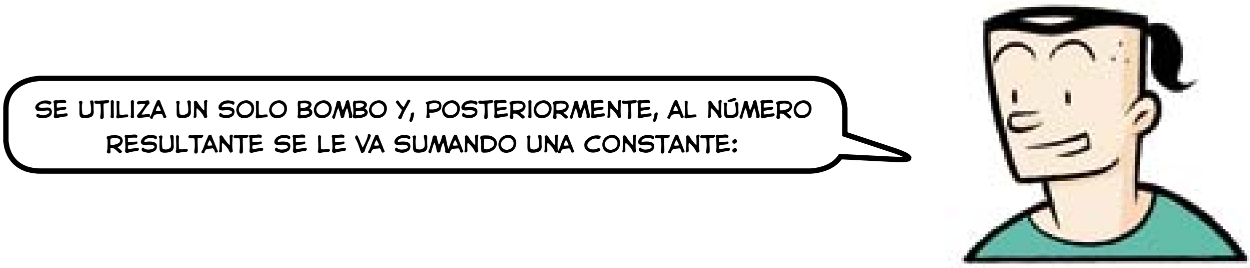
3690	2492	7171	7720	6509	7549	2830	5733	4780
0813	6790	6858	1489	2669	8743	1901	4971	8280
6177	5289	1002	4223	6451	7632	7577	2316	9002
0772	2160	7236	0812	4195	5509	0830	8261	9232
5692	9870	3583	8997	1535	5466	8830	7271	3809
2030	3828	7880	0586	8482	7811	6807	3509	2729
1039	3382	7600	1077	4455	8806	1822	1669	7501
7227	0104	4141	1521	9104	5563	1392	8238	4882
8506	6348	[illegible]	8252	1062	1757	0964	2983	2244
5086	0303	7423	3298	3975	2831	2237	1508	7642
0192	1629	0577	3590	2209	4839	6332	1490	3092
0935	5565	2315	8030	7651	5189	0075	9353	1921
2505	3073	0201	4113	2677	0394	8601	3310	8389
7277	9889	0390	3579	4620	5550	0210	2082	4664
5484	3900	3485	0741	9069	5920	4326	7704	6525
0905	7127	5933	1137	7583	0450	5058	7978	3444

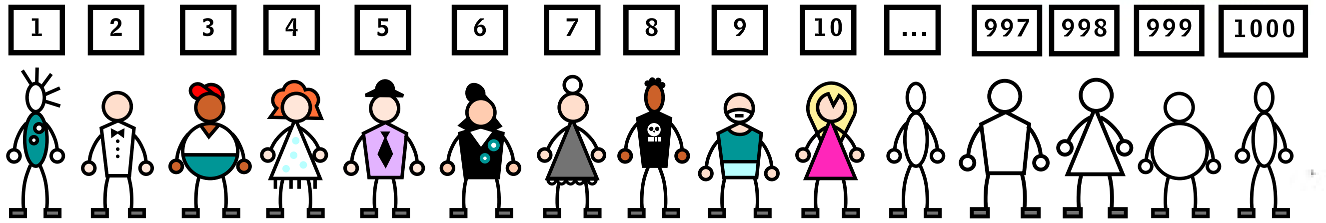

TAMBIÉN SERÍA EXTRAORDINARIO TOMAR COMO LISTA DE NÚMEROS ALEATORIOS LA DE LOS PRIMEROS PREMIOS DE LA LOTERÍA DURANTE MUCHOS AÑOS.
NO ES NADA EXTRAORDINARIO, PORQUE SON NÚMEROS ALEATORIOS.


NO, SI LO DE EXTRAORDINARIO LO DIGO, PORQUE SI DEMOSTRAMOS QUE NO LO SON... ¡QUÉ "MONTÓN" DE DINERO ÍBAMOS A GANAR!
CREO QUE DEBERÍAMOS EMPEZAR CON EL MUESTREO ALEATORIO SISTEMÁTICO.


PUES ESCENIFIQUÉMOSLO:
HAN SALIDO: DEL PRIMER BOMBO, EL 5...
Y DEL SEGUNDO, UN 7.
5
7


EMPECEMOS EL PROCESO:
1 2 3 4 5 6 7 8 9 10 11 12 13 14 ...
5 +7 12 +7 19 +7 26 +7 33 ...
Y ASÍ SEGUIRÍAMOS CON LA MUESTRA HASTA QUE EL NÚMERO DE SELECCIONADOS FUERA IGUAL A LA CANTIDAD MUESTRAL.
EL MUESTREO ALEATORIO SISTEMÁTICO ES COMO EL MUESTREO ALEATORIO SIMPLE, PERO CON UN SISTEMA PREVIO DE REALIZACIÓN; EL QUE SE HA EXPUESTO ANTES O...





Población (tamaño $= \text{N} = 1000$)

Tamaño muestral $= \text{n} = 50$

$$\text{Constante} = \text{k} = \frac{N}{n} = \frac{1000}{50} = 20$$

Sorteo $\longrightarrow$ 4

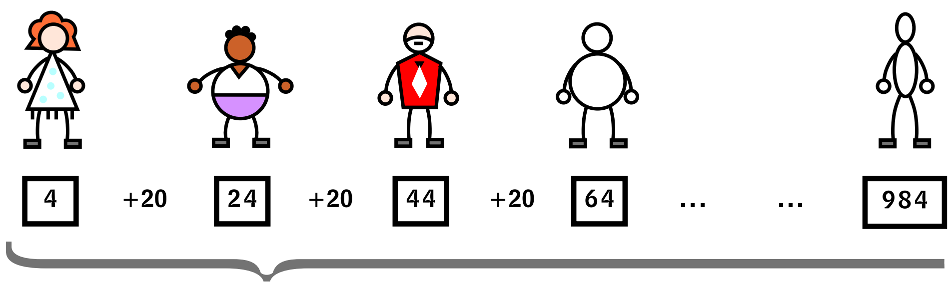
1º de la muestra $= 4$
2º $= 4 + 20 = 24$
3º $= 4 + 2 \times 20 = 44$
4º $= 4 + 3 \times 20 = 64$

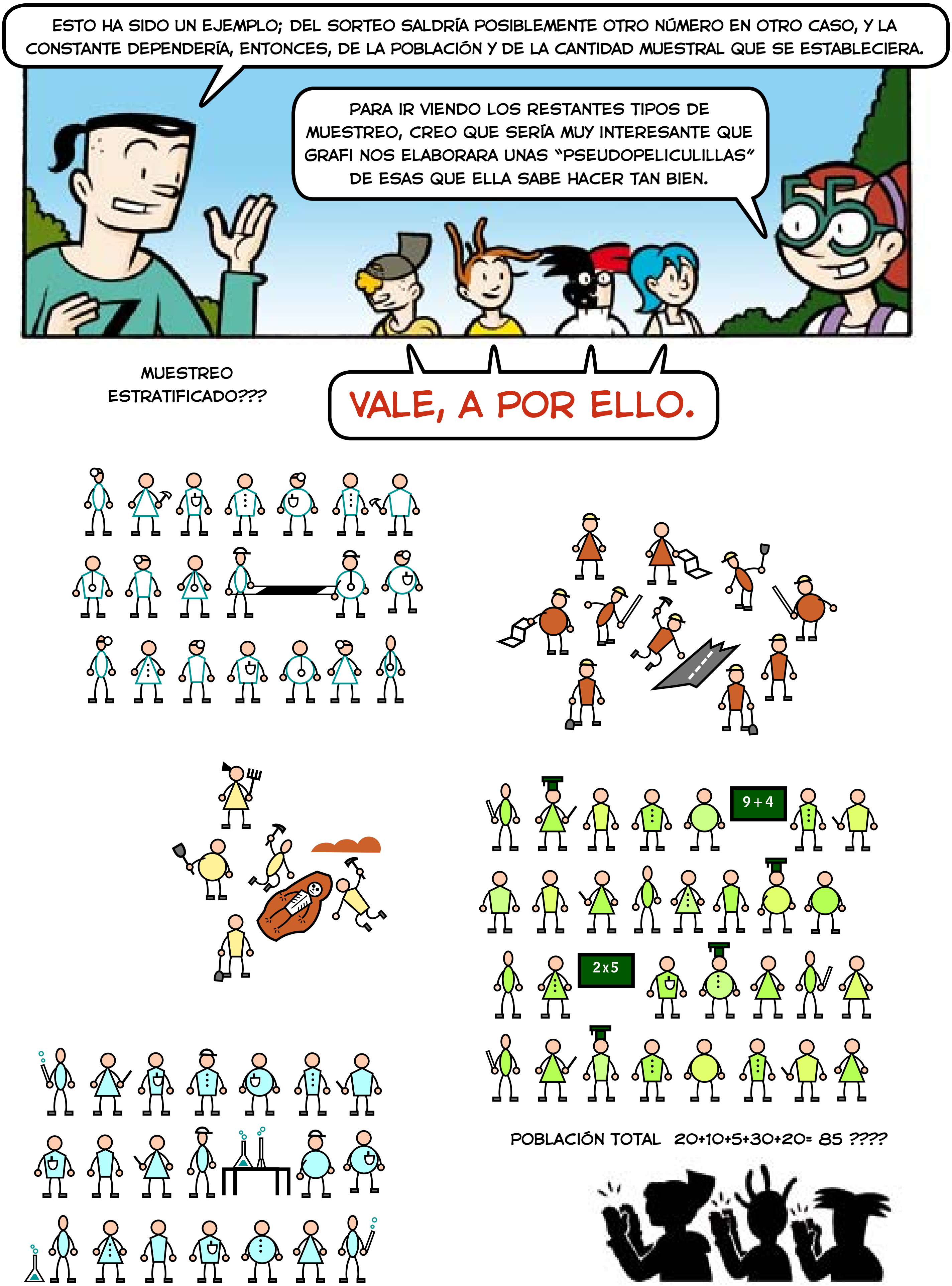
. . . .

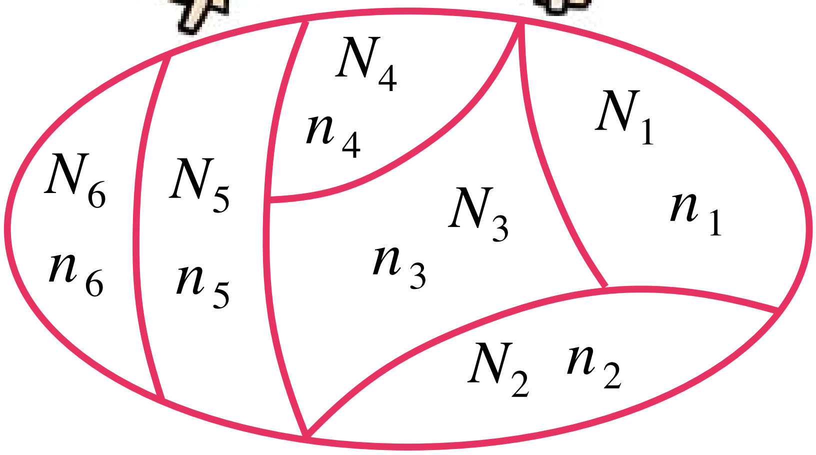
25º $= 4 + 24 \times 20 = 484$

. . . .

50º $= 4 + 49 \times 20 = 984$




ESTO HA SIDO UN EJEMPLO; DEL SORTEO SALDRÍA POSIBLEMENTE OTRO NÚMERO EN OTRO CASO, Y LA CONSTANTE DEPENDERÍA, ENTONCES, DE LA POBLACIÓN Y DE LA CANTIDAD MUESTRAL QUE SE ESTABLECIERA.
PARA IR VIENDO LOS RESTANTES TIPOS DE MUESTREO, CREO QUE SERÍA MUY INTERESANTE QUE GRAFI NOS ELABORARA UNAS "PSEUDOPELICULILLAS" DE ESAS QUE ELLA SABE HACER TAN BIEN.
MUESTREO ESTRATIFICADO???
VALE, A POR ELLO.
9 + 4
2 x 5
POBLACIÓN TOTAL 20+10+5+30+20= 85 ????



Tamaños poblacionales en cada estrato:

$N_1; N_2; N_3; N_4; N_5; N_6$

Tamaño poblacional:

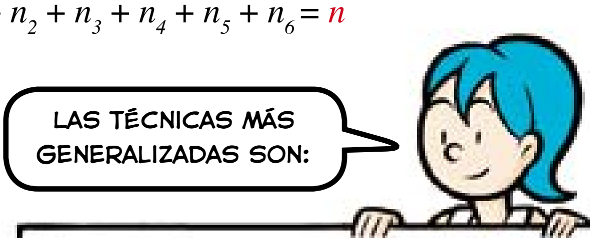
$N_1 + N_2 + N_3 + N_4 + N_5 + N_6 = N$

Tamaños muestrales en cada estrato:

$n_1; n_2; n_3; n_4; n_5; n_6$

Tamaño muestral:

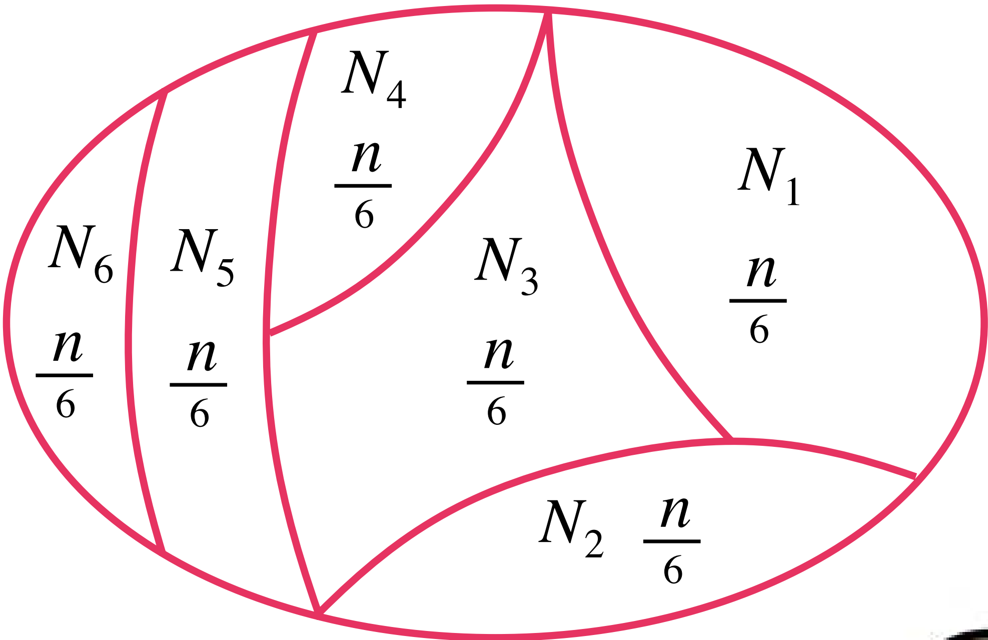
$n_1 + n_2 + n_3 + n_4 + n_5 + n_6 = n$



1°. Muestreo estratificado proporcional.

2°. Muestreo estratificado no proporcional.

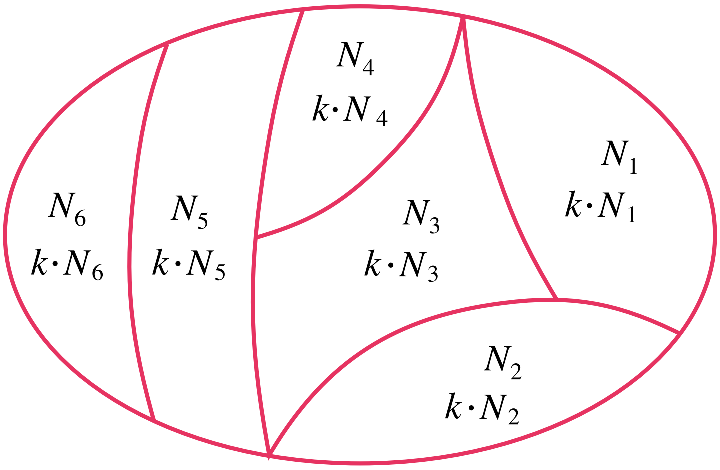
3°. Asignación óptima de los estratos.



Afijación uniforme

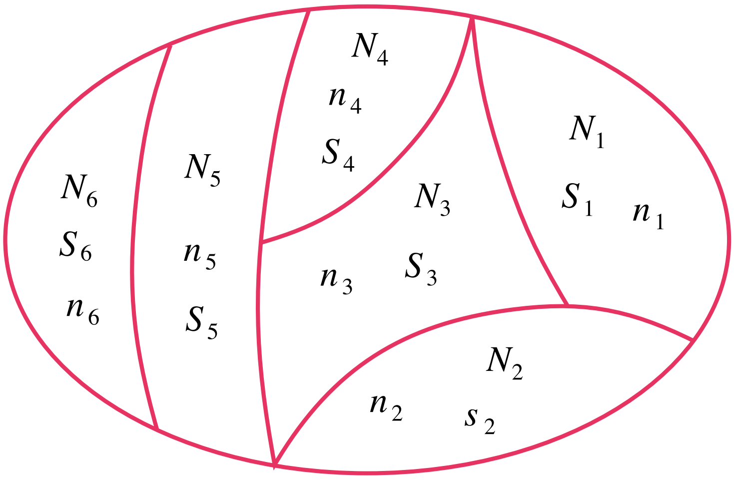
Unidades muestrales de cada estrato:

$$n_1 = n_2 = n_3 = n_4 = n_5 = n_6 = \frac{n}{6}$$



$$\frac{n}{N} = k$$

$$n_1 = k \times N_1 \qquad n_4 = k \times N_4$$
$$n_2 = k \times N_2 \qquad n_5 = k \times N_5$$
$$n_3 = k \times N_3 \qquad n_6 = k \times N_6$$



Desviaciones típicas o estándares de cada estrato:

$S_1; S_2; S_3; S_4; S_5; S_6$

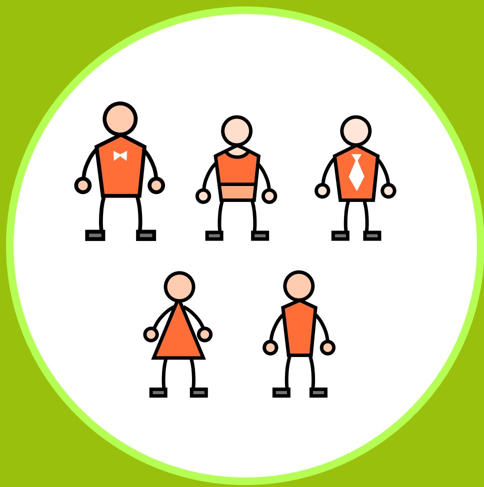
Tamaños muestrales de cada estrato:

$$n_1 = n \times \frac{N_1 S_1}{\sum_{i=1}^{6} N_i S_i} \qquad n_4 = n \times \frac{N_4 S_4}{\sum_{i=1}^{6} N_i S_i}$$

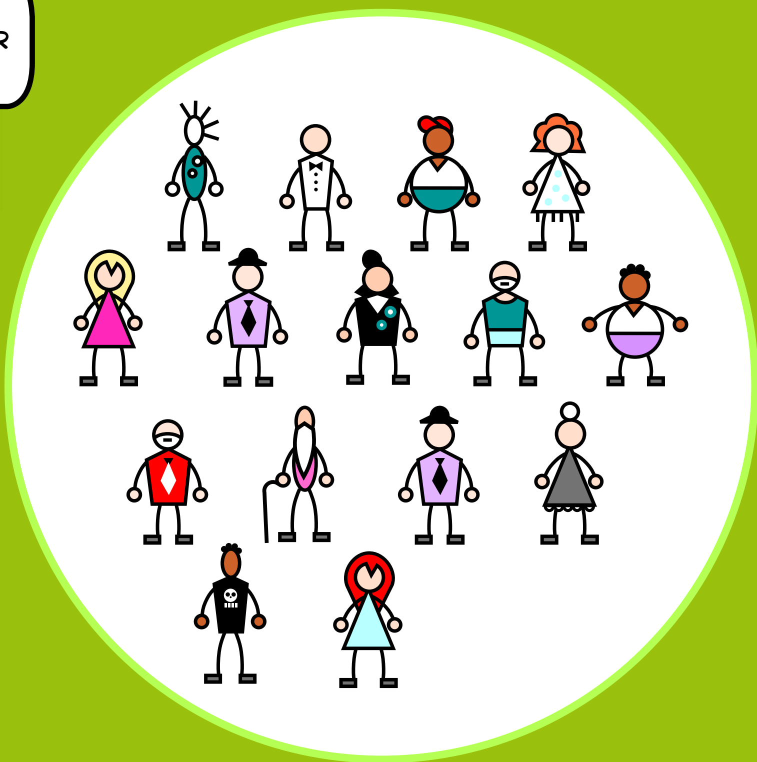
$$n_2 = n \times \frac{N_2 S_2}{\sum_{i=1}^{6} N_i S_i} \qquad n_5 = n \times \frac{N_5 S_5}{\sum_{i=1}^{6} N_i S_i}$$

$$n_3 = n \times \frac{N_3 S_3}{\sum_{i=1}^{6} N_i S_i} \qquad n_6 = n \times \frac{N_6 S_6}{\sum_{i=1}^{6} N_i S_i}$$

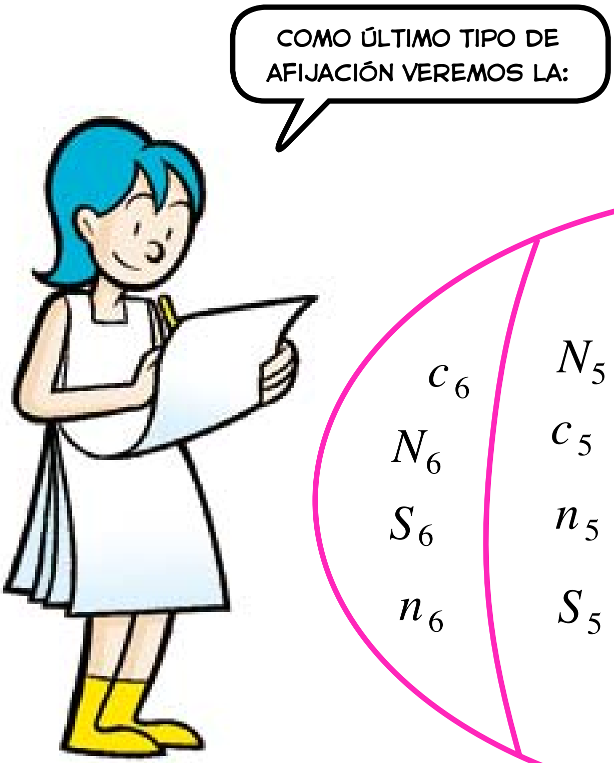
Estrato 1



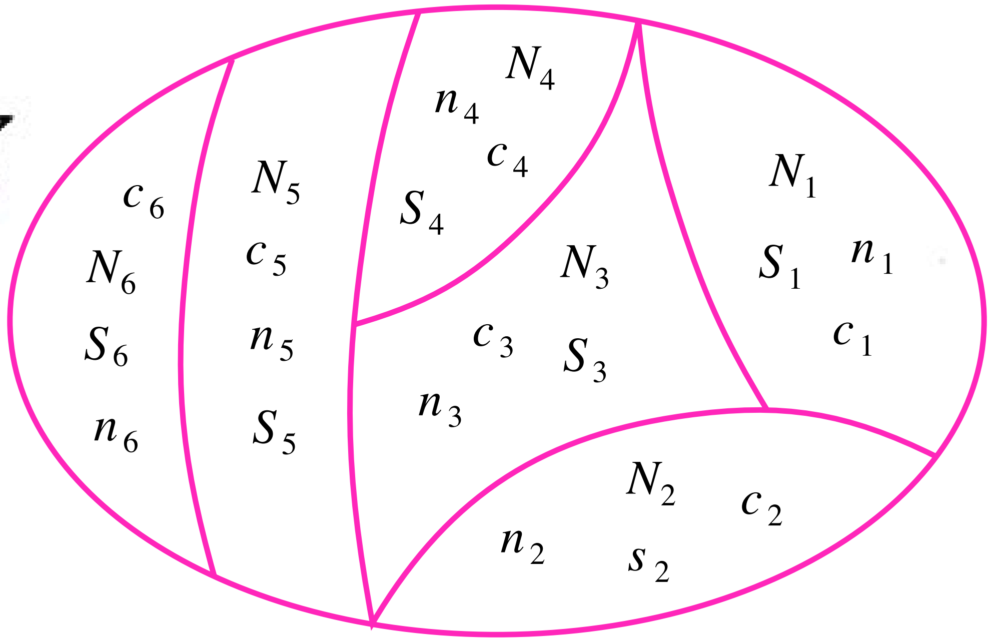
Estrato 3




Estrato 2
Estrato 4
Estrato 3: Tamaño "grande" Variabilidad "grande"
Estrato 4: Tamaño "pequeño" Variabilidad "grande"
Estrato 2: Tamaño "grande" Variabilidad "pequeña"
Estrato 1: Tamaño "pequeño" Variabilidad "pequeña"
EL TAMAÑO MUESTRAL TIENE ESE MISMO ORDEN: 3, 4, 2, 1 DE MAYOR A MENOR.



Afijación óptima

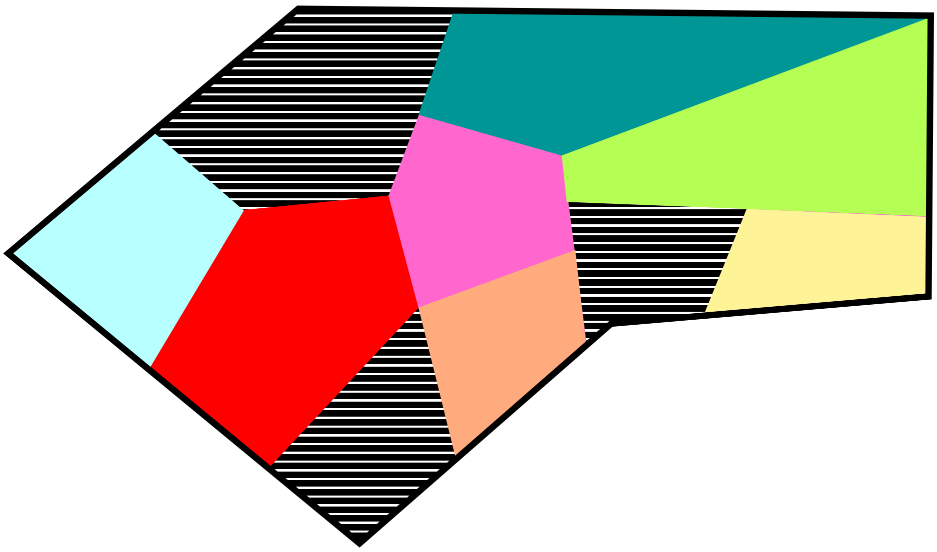


En este tipo se tiene también en cuenta el coste. Sus fórmulas son:

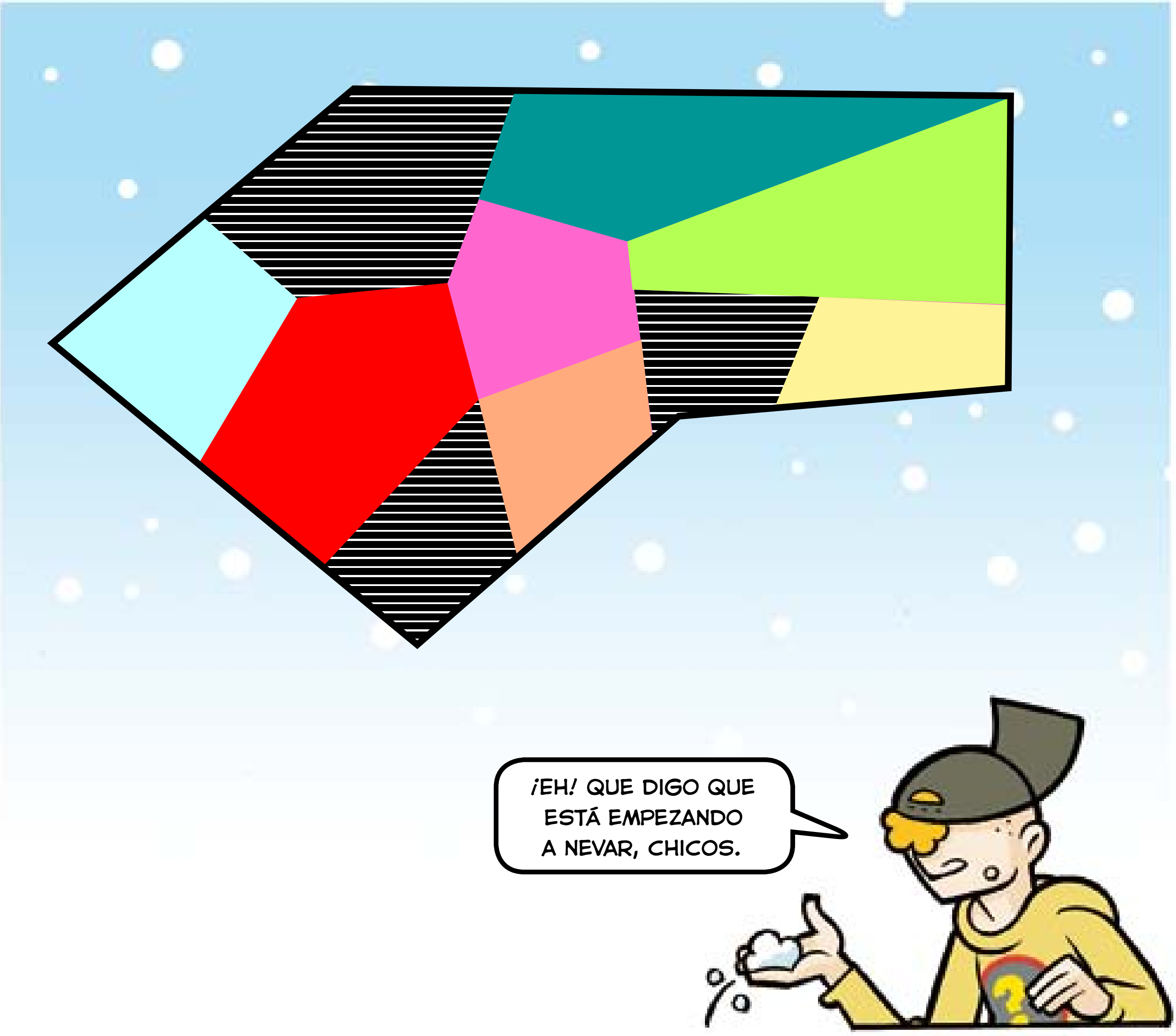
$$n_1 = n \times \frac{\dfrac{N_1 S_1}{\sqrt{c_1}}}{\displaystyle\sum_{i=1}^{6} \dfrac{N_i S_i}{\sqrt{c_i}}} \qquad n_4 = n \times \frac{\dfrac{N_4 S_4}{\sqrt{c_4}}}{\displaystyle\sum_{i=1}^{6} \dfrac{N_i S_i}{\sqrt{c_i}}}$$

$$n_2 = n \times \frac{\frac{N_2 S_2}{\sqrt{c_2}}}{\sum_{i=1}^{6} \frac{N_i S_i}{\sqrt{c_i}}} \qquad n_5 = n \times \frac{\frac{N_5 S_5}{\sqrt{c_5}}}{\sum_{i=1}^{6} \frac{N_i S_i}{\sqrt{c_i}}}$$

$$n_3 = n \times \frac{\frac{N_3 S_3}{\sqrt{c_3}}}{\sum_{i=1}^{6} \frac{N_i S_i}{\sqrt{c_i}}} \qquad n_6 = n \times \frac{\frac{N_6 S_6}{\sqrt{c_6}}}{\sum_{i=1}^{6} \frac{N_i S_i}{\sqrt{c_i}}}$$




PRIMERO, DESAGREGAREMOS LA POBLACIÓN EN GRUPOS HETEROGÉNEOS DE INDIVIDUOS.
PARECE QUE ESTÁ CAMBIANDO EL TIEMPO...
SELECCIONAMOS AL AZAR UNA MUESTRA DE ESTOS GRUPOS, TAMBIÉN DENOMINADOS ÁREAS O CONGLOMERADOS.


¡EH! QUE DIGO QUE ESTÁ EMPEZANDO A NEVAR, CHICOS.


Y DE LOS CONGLOMERADOS SELECCIONADOS COMO MUESTRA, HACEMOS UN ESTUDIO CENSAL, O SEA, ESTUDIAMOS A TODOS LOS INDIVIDUOS QUE LOS COMPONEN.
Y ASÍ PODRÍAMOS CONTINUAR... CON MUCHAS FORMAS DE MUESTREO.
LO QUE TENEMOS QUE HACER ES CONTINUAR A CUBIERTO.

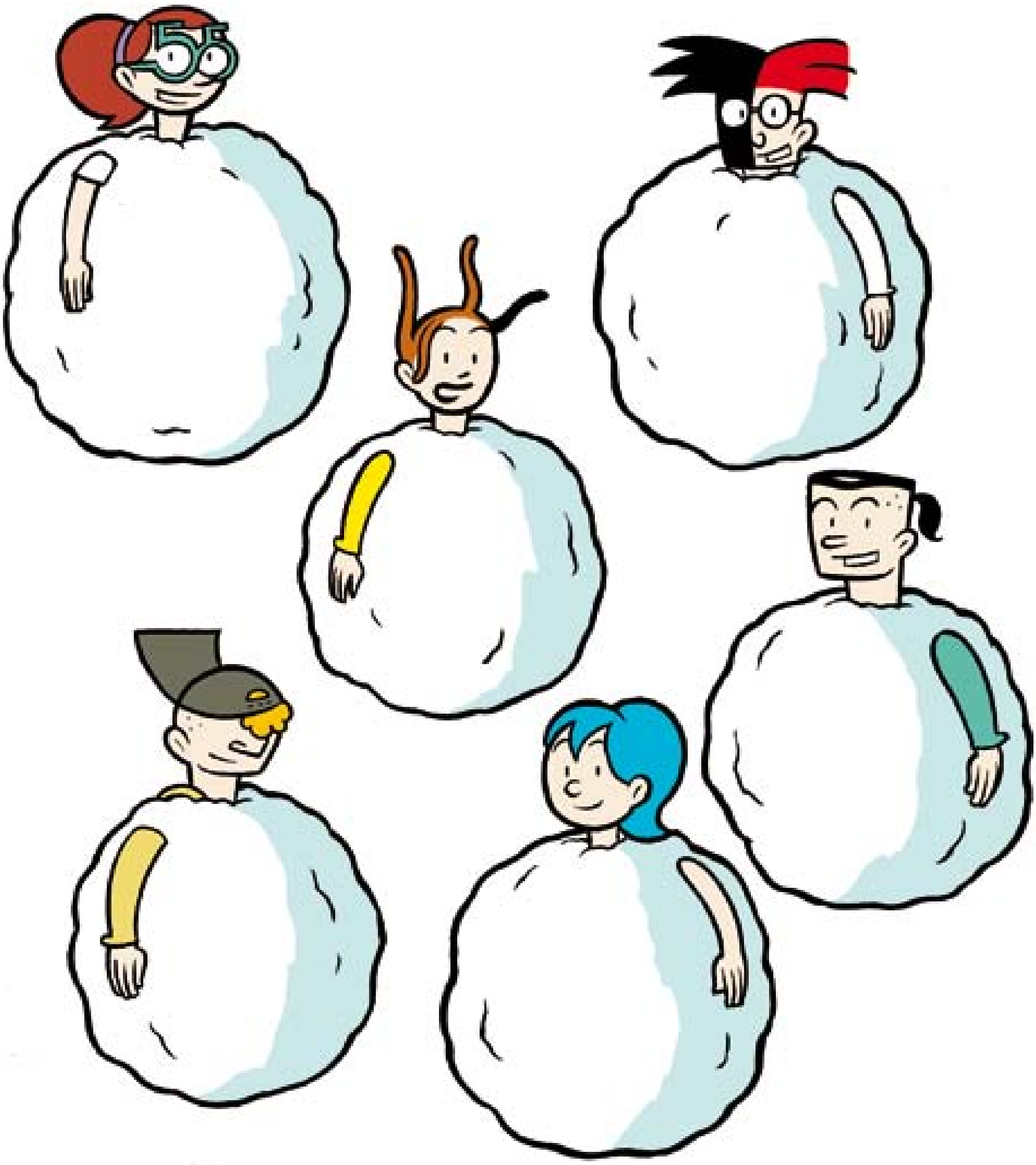
Proceso del muestreo

- Definición de los objetivos de la encuesta.
- Definición de la población objeto de estudio.
  Elementos. Unidades de muestreo.
  Alcance. Tiempo.
- Definición del marco muestral.
- Selección del procedimiento de muestreo.
- Establecimiento de la medida de la muestra.
- Obtención de la muestra.


¡QUÉ NEVADA!
ANTES DE QUE NOS MARCHEMOS... HE ENCONTRADO UN TIPO DE MUESTREO QUE LE VA COMO ANILLO AL DEDO.
¿CÓMO DICES?
MUESTREO EN BOLA DE NIEVE: CONSISTE EN HACER UNA PRIMERA SELECCIÓN Y QUE LOS ENCUESTADOS TE RECOMIENDEN A OTROS PARA ENCUESTAR Y ASÍ...


LAS COSAS ESTÁN CAMBIANDO... ANTES ACERTIJO SIEMPRE ERA EL PRIMERO EN QUERERSE IR, Y HOY ES EL ÚLTIMO INVESTIGANDO.


¡¡¡A LA NIEVE!!!



FIN

GOVERN de les ILLES BALEARS