

Estadística – Licenciatura en Nutrición - UCEL

Capítulo 3: Resumen de datos en forma gráfica

RESOLUCIÓN DE EJERCICIOS PROPUESTOS EN LAS DIAPOSITIVAS

Manos a la Obra N° 1

¿QUÉ TIPO DE VARIABLE ES?

- 1) Variable cualitativa. Algunos valores posibles: Kellogg's, Granix, Quaker.
- 2) Variable cuantitativa continua. Algunos valores posibles: 2° C, 10° C.
- 3) Variable cuantitativa continua. Algunos valores posibles: 124 minutos, 1 hora y media.
- 4) Variable cuantitativa discreta. Algunos valores posibles: 1 hora, 3 horas.
- 5) Variable cuantitativa continua. Algunos valores posibles: 200 ml, 0,25 litros.
- 6) Variable cuantitativa discreta. Algunos valores posibles: 15, 29, 35. Nota: El valor máximo para esta variable es 35.
- 7) Variable cualitativa. Los únicos valores posibles de la variable son los que indica la consigna.

Ejercicio propuesto (diapositiva 15)

Segunda Fila:

- Hay 30 personas con peso normal.
- La proporción de personas de peso normal es 0,60.
- El 60% de las personas tienen peso normal.

Tercera Fila:

- 11 de las personas estudiadas presentan sobrepeso.
- La proporción de personas estudiadas que presentan sobrepeso es 0,22.
- El 22% de las personas estudiadas tienen sobrepeso.

Cuarta Fila:

- Hay 2 personas obesas en la muestra.
- La proporción de personas obesas en la muestra es 0,04.
- El 4% de las personas en la muestra son obesas.

Manos a la Obra N° 2

PARA PENSAR PÁGINA 144

En el gráfico de barras vemos representados de forma simultánea los resultados de los dos meses estudiados: enero de 1992 (en blanco) y enero de 1996 (rayado). La variable estudiada es

el problema más importante que enfrenta américa (Estados Unidos de América) en opinión de quien responde; esta es una variable cualitativa, y los valores observados en el gráfico son: déficit, crimen/violencia, economía, desempleo, falta de vivienda, drogas, educación e inflación. Entre las categorías de la variable no se puede establecer ningún orden. Tengamos en cuenta, también, que en este gráfico de barras están representados los porcentajes de cada categoría de la variable en estudio.

- a) Para responder esta pregunta, primero debemos comparar el ancho de las barras blancas, que corresponden a las respuestas de 1992. La barra más ancha es la que corresponde a la economía, es decir que la categoría que recibió el mayor porcentaje de respuestas fue que el problema más importante que enfrenta américa es la economía. Es más, podemos decir que aproximadamente el 30% de los encuestados en enero de 1992 considera que el problema más serio que enfrenta américa es la economía. Para responder a la segunda parte de la pregunta, debemos comparar el ancho de las barras que corresponden a enero de 1996. En este mes, la barra más ancha es la que corresponde a déficit, es decir que el problema que recibió el mayor porcentaje de respuestas en enero de 1992 y en enero de 1996 no es el mismo.
- b) En enero de 1992, aproximadamente el 1% de los encuestados respondió que “crimen y violencia” es el problema más importante que enfrenta américa (ver el ancho de la barra blanca). En enero de 1992, aproximadamente el 14% de los entrevistados respondió que el problema más importante que enfrenta américa es “crimen y violencia” (ancho de la barra rayada).
- c) La suma de los porcentajes de las barras blancas, correspondientes a enero de 1992, es de aproximadamente 70% (en orden: $5+1+30+20+5+5+1+1=68\%$). En los gráficos de barras, la suma de las frecuencias absolutas o relativas, o porcentajes debería ser igual a n , 1 o 100%, respectivamente. Una explicación posible en este caso es que haya más respuestas posibles a la pregunta, es decir, más valores posibles de la variable (por ejemplo pobreza, salud, desnutrición) pero que estos valores no hayan sido representados en el gráfico porque presentan porcentajes de ocurrencia muy bajos.

PARA RESOLVER 3.6 PÁGINA 144

Aclaración: Polimodal fue un tipo de educación usado en Argentina, que reemplazaba a la escuela secundaria. Estaba formada por los últimos tres años de educación escolar, y una vez aprobado este nivel el alumno podía acceder a la educación universitaria.

La resolución está desarrollada en la actividad.

EJERCICIO 2 DE LA PRÁCTICA ADICIONAL

- 1) La variable estudiada es “calificación que recibe el seminario”. Esta es una variable cualitativa. Los números sólo son un código para simplificar el registro de los datos, pero no tienen significado como tales. Los valores posibles de la variable son: mediocre, aceptable, bueno, muy bueno y excelente. Además, existe un orden entre las categorías de la variable, por ejemplo, la opinión de que el seminario fue “bueno” es mejor que las opiniones que dicen que el seminario fue “aceptable” o “mediocre”.

- 2) Para resumir una variable de tipo cualitativa pueden usarse el gráfico de sectores circulares o el gráfico de barras. Lo primero que hay que hacer para poder construir cualquiera de estos dos gráficos es armar la tabla de frecuencias. En las filas vamos a indicar las categorías de la variable. Indicamos entre paréntesis el código que corresponde a cada clase sólo para recordarlo, pero no debemos incluirlo si hacemos una presentación de los resultados:

Categoría (código)	Frecuencia Absoluta	Frecuencia Relativa	Frecuencia Relativa Porcentual (%)
Mediocre (1)			
Aceptable (2)			
Bueno (3)			
Muy Bueno (4)			
Excelente (5)			

Para completar la tabla, primero contamos cuántos médicos dieron cada respuesta. Por ejemplo, ¿cuántos médicos pensaron que el seminario fue mediocre? Volvemos a la tabla de datos y contamos cuántas respuestas iguales a 1 (mediocre) hay:

3	4	4	5	1	5	3	4	5	2
4	5	3	4	4	4	5	5	4	1
4	5	4	2	5	4	2	4	4	4
5	5	3	4	5	5	2	4	3	4
5	4	3	5	4	4	3	5	4	5
4	3	5	3	4	4	3	5	3	3

Hay dos médicos que indicaron que el seminario les pareció mediocre. Así podemos contar cuántas respuestas hay de cada categoría de la variable.

Categoría	Frecuencia Absoluta	Frecuencia Relativa	Frecuencia Relativa Porcentual (%)
Mediocre	2		
Aceptable	4		
Bueno	12		
Muy Bueno	24		
Excelente	18		

Consejo útil: revisar que la suma de las frecuencias absolutas sea igual al tamaño de muestra n . En este caso, la suma de las frecuencias absolutas es $n=60$.

Para calcular las frecuencias relativas, dividimos cada frecuencia absoluta por el tamaño de muestra. Para obtener la frecuencia relativa porcentual, multiplicamos la frecuencia relativa por 100.

Categoría	Frecuencia Absoluta	Frecuencia Relativa	Frecuencia Relativa Porcentual (%)
Mediocre	2	$2/60 = 0,03$	$0,03 \times 100 = 3$
Aceptable	4	$4/60 = 0,07$	$0,07 \times 100 = 7$
Bueno	12	$12/60 = 0,20$	$0,20 \times 100 = 20$
Muy Bueno	24	$24/60 = 0,40$	$0,40 \times 100 = 40$
Excelente	18	$18/60 = 0,30$	$0,30 \times 100 = 30$

Nota importante I: la suma de las frecuencias relativas debe ser igual a 1. LA suma de las frecuencias relativas porcentuales debe ser igual a 100.

Nota importante II: al calcular las frecuencias relativas, solemos presentar los resultados con 2 o 3 decimales. Cuando el resultado tiene más decimales, redondeamos las frecuencias decimales de tal manera que la suma de todas sea igual a 1.

Pasando los resultados en limpio:

Categoría	Frecuencia Absoluta	Frecuencia Relativa	Frecuencia Relativa Porcentual (%)
Mediocre	2	0,03	3
Aceptable	4	0,07	7
Bueno	12	0,20	20
Muy Bueno	24	0,40	40
Excelente	18	0,30	30

Con esta tabla de frecuencias podemos construir tanto el gráfico de sectores circulares con el gráfico de barras.

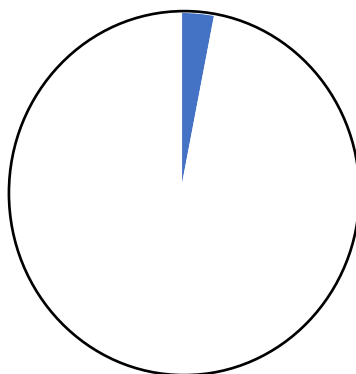
Gráfico de sectores circulares

Para construir este gráfico, tomamos la frecuencia absoluta (o relativa, o relativa porcentual, es lo mismo) y calculamos qué parte de la circunferencia representa cada categoría. Podemos hacer una regla de tres simple considerando que la circunferencia tiene 360° , o multiplicar las frecuencias relativas por 360° :

Categoría	Frecuencia Relativa	Ángulo ($^\circ$)
Mediocre	0,03	$0,03 \times 360 = 11$
Aceptable	0,07	$0,07 \times 360 = 25$
Bueno	0,20	$0,20 \times 360 = 72$
Muy Bueno	0,40	$0,40 \times 360 = 144$
Excelente	0,30	$0,30 \times 360 = 108$

Sólo tenemos que tener cuidado de que la suma de los ángulos sea igual a 360. Estos son los ángulos que forman los segmentos que definen cada sector circular. Para empezar la representación, ubicamos el comienzo del sector que corresponde a la primera categoría en 90° . Vamos a respetar el orden entre las categorías porque en esta variable tiene sentido. La categoría “mediocre” queda representada:

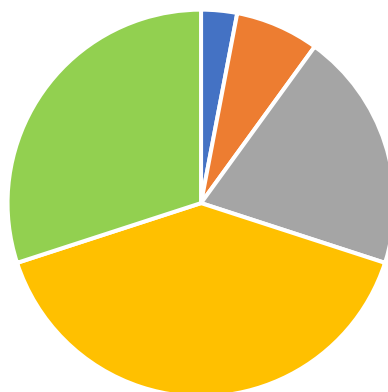
Distribución de la opinión sobre el seminario



■ Mediocre Aceptable Bueno Muy Bueno Excelente

La amplitud de este ángulo es de 11° . Agregamos sucesivamente los demás sectores circulares y obtenemos:

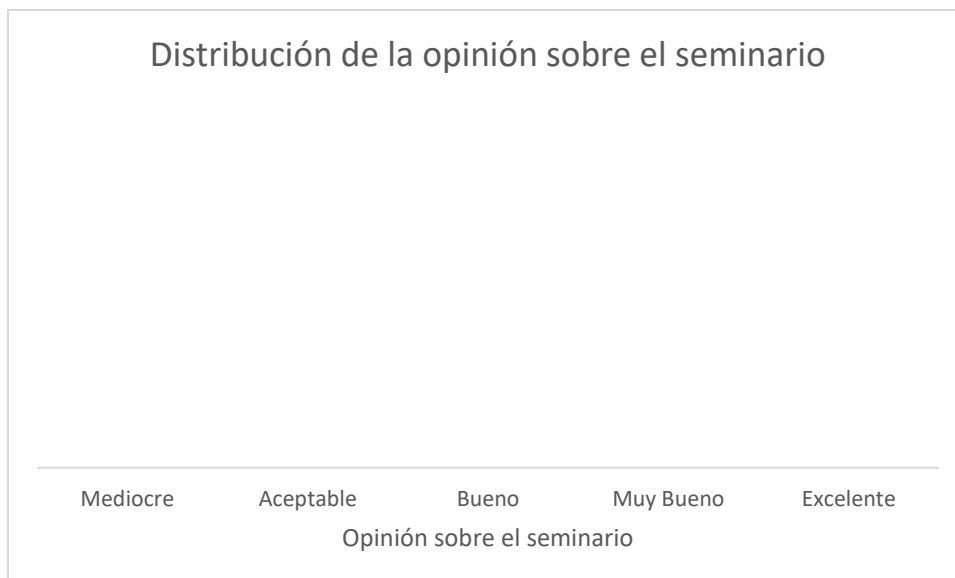
Distribución de la opinión sobre el seminario



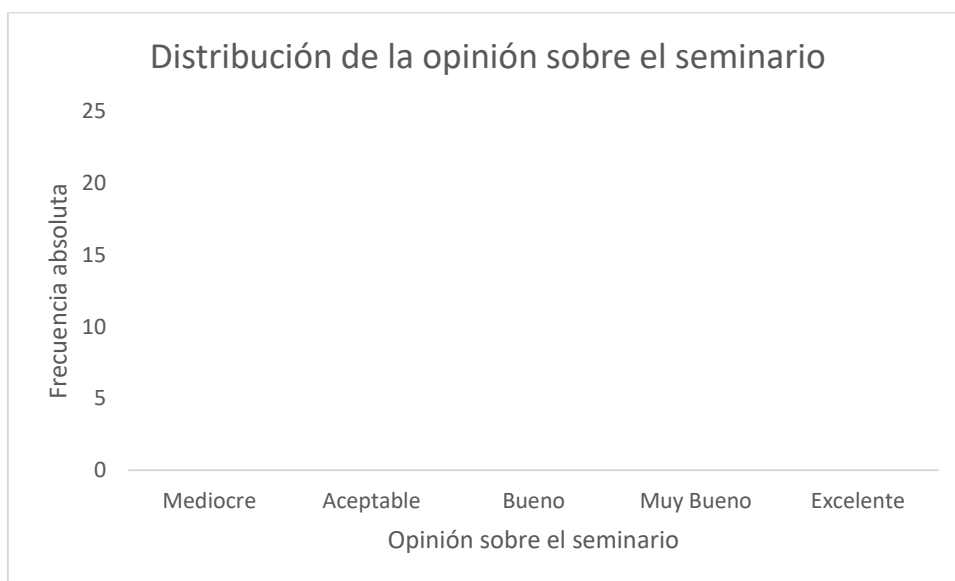
■ Mediocre ■ Aceptable ■ Bueno ■ Muy Bueno ■ Excelente

Gráfico de barras

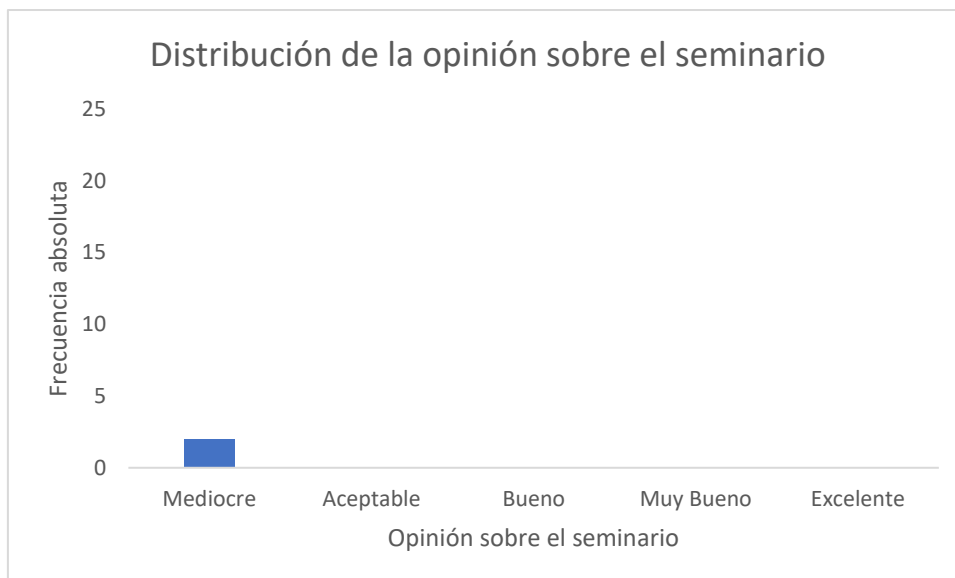
Para construir el gráfico de barras con las frecuencias absolutas, primero construimos el eje x, colocando las categorías de la variable de manera equiespaciada (la misma distancia entre cada par de categorías):



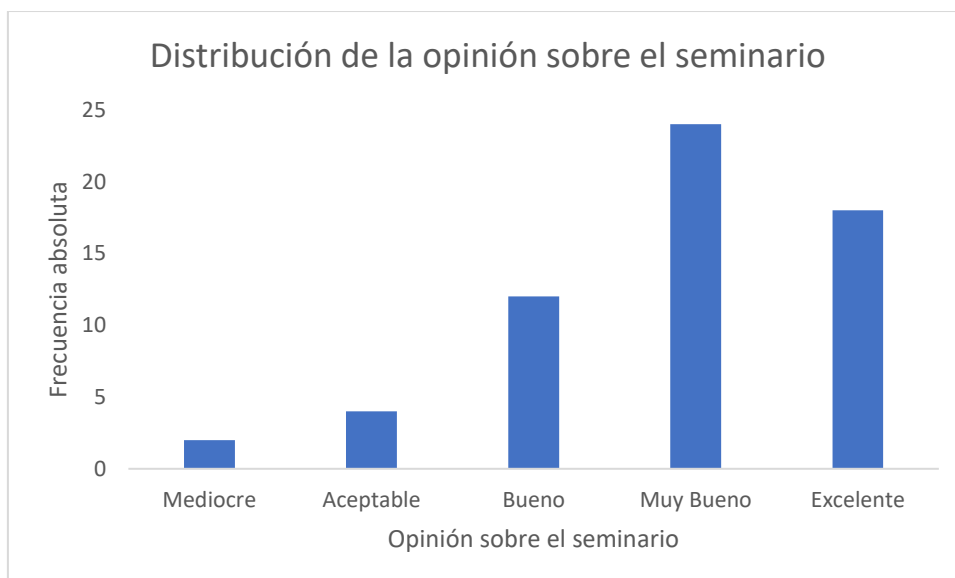
Luego, agregamos los valores en el eje y. Vamos a empezar el eje con el valor 0 y terminarlo cerca de la frecuencia absoluta más grande, que es 24. Por ejemplo, podríamos terminarlo en 25.



Luego, para cada categoría elevamos una barra desde el eje x hasta la frecuencia absoluta. Todas las barras deben tener el mismo ancho. Para la respuesta de que el seminario fue “mediocre”, la barra llega hasta 2:



El gráfico completo resulta:



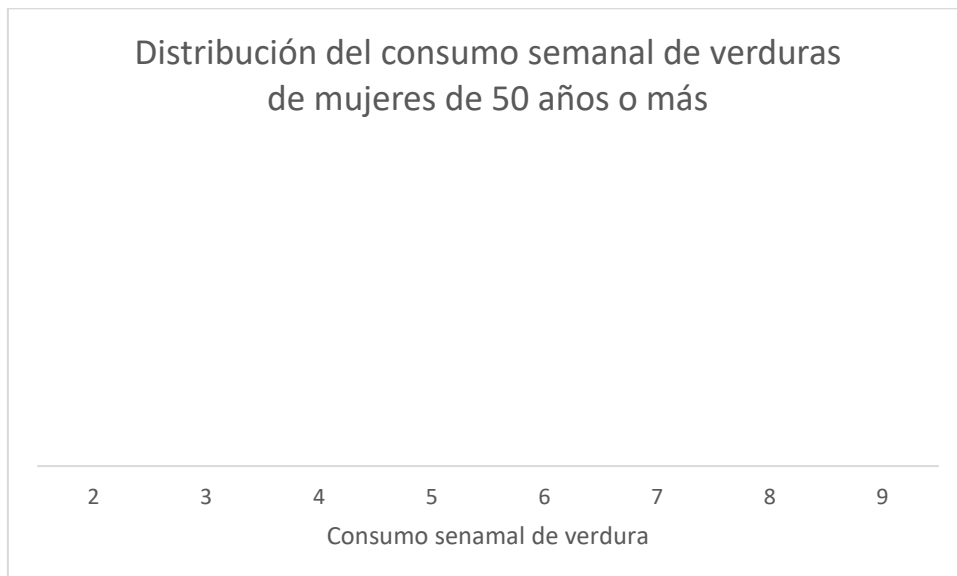
Nota importante: ¡No debemos olvidar los nombres de los ejes y el título del gráfico!

Manos a la Obra N° 3

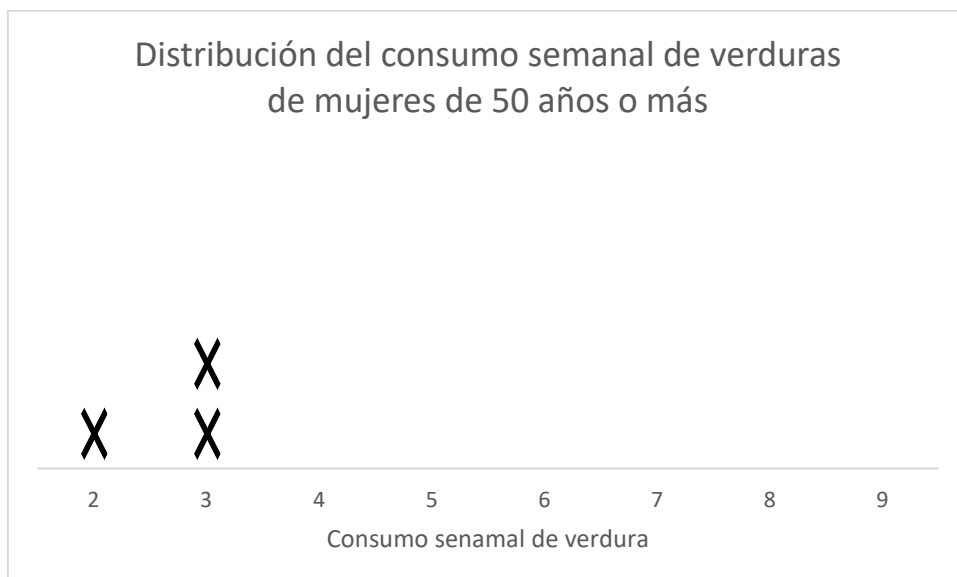
La variable en estudio es el consumo semanal de verduras de mujeres de 50 años o más. Esta es una variable cuantitativa discreta. Los consumos semanales de verduras de las mujeres en la muestra ordenados son:

2	3	3	6	6	7	7	7	8	8	9	9	9	9
---	---	---	---	---	---	---	---	---	---	---	---	---	---

Para construir el gráfico de frecuencias, ubicamos el mínimo consumo semanal de verduras (2) y el máximo consumo semanal de verduras (9) en el eje x y marcamos todos los valores entre estos dos números respetando una escala:



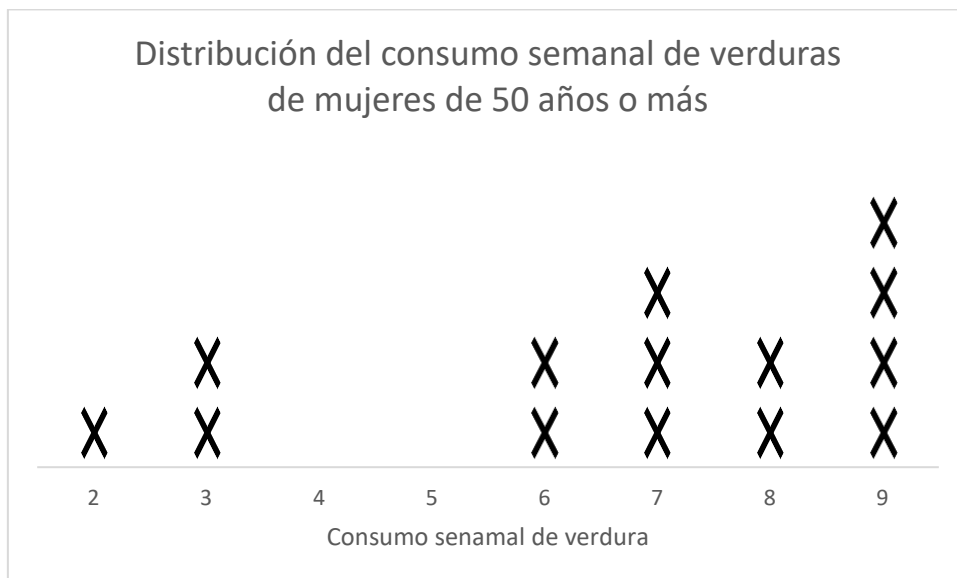
Luego, tenemos que marcar con una cruz el consumo semanal de verduras de cada una de las mujeres de 50 años o más en el gráfico. Por ejemplo, hay una mujer que consume 2 porciones de verduras por semana, mientras que dos mujeres consumen 3 porciones semanales de verdura. Eso queda representado así:



También podemos armar una tabla de frecuencias antes de construir el gráfico. La frecuencia absoluta es el número de cruces.

Consumo de verduras	Frecuencia absoluta
2	1
3	2
6	2
7	3
8	2
9	4

El gráfico completo es:



- 1) Las mujeres que consumen 2 o 3 porciones semanales de verdura pueden considerarse outliers.
- 2) Esas mismas mujeres forman un conglomerado.
- 3) Hay una brecha entre el consumo de semanal de 3 porciones de verdura y el consumo semanal de 6 porciones de verdura.
- 4) El consumo semanal de verduras de las mujeres de 50 años o más varía entre 2 y 9 porciones por semana. La mayoría de las mujeres consume entre 6 y 9 porciones semanales de verdura. La distribución es asimétrica hacia la izquierda.

Manos a la Obra N° 4

Para armar el diagrama de tallo y hoja de la altura de los hombres que tienen peso normal, podemos ordenar primero la altura de menor a mayor:

1,57	1,66	1,67	1,67	1,72	1,73	1,74	1,75
1,75	1,77	1,79	1,80	1,81	1,84	1,86	1,92

Dejamos como hojas el último dígito del número. Por ejemplo, para el hombre que mide 1,57 m, el tallo es 1,5 y la hoja es 7. El tallo 1,6 tiene tres hojas: 6 (del hombre que mide 1,66), 7 y 7 (de los dos hombres que miden 1,67). El diagrama de tallo y hoja queda:

1,5		7
1,6		6 7 7
1,7		2 3 4 5 5 7 9
1,8		0 1 4 6
1,9		2

Nota: 1,5|7 se lee 1,57 m

Tenemos que recordar la nota al final del diagrama para aclarar su lectura.

- 1) En este diagrama no se observan outliers.
- 2) La distribución de la altura de los hombres es aproximadamente simétrica.
- 3) Hay 10 hombres con peso normal en la muestra. Necesitamos obtener dos subconjuntos de 8 hombres:

1,5		7
1,6		6 7 7
1,7		2 3 4 5 5 7 9
1,8		0 1 4 6
1,9		2

Como tenemos dos valores iguales, podemos decir que el 50% de los hombres mide 1,75 m o menos, mientras que el otro 50% de los hombres mide 1,75 m o más.

PARA PENSAR DE LA PÁGINA 163

El diagrama N°1 no representa todos los tallos. Por ejemplo, pasa del tallo 27 al 32, cuando entre esos dos valores deberían estar los tallos 29, 29, 30 y 31.

El diagrama N°2 está demasiado concentrado, podría subdividirse cada tallo en dos partes. Por ejemplo, tener un tallo 2 que tenga las hojas desde el 0 hasta el 4, y un tallo 2* que tenga las hojas desde 5 hasta 9.

El diagrama N°3 tiene demasiados tallos y pocas hojas. Debería elegirse alguna otra representación gráfica para esta variable.

EJERCICIO 3.13 DE LA PÁGINA 164

A continuación se presentan los cuatro diagramas de tallo y hoja posibles:

Diagrama de tallo y hoja del contenido de grasa en 3 onzas de pescado.

0	5	8	9
1	3	5	7
2	5		
3	8		
4	2	4	
5	3		
6	9		
7			
8			
9			
10			
11	4		

Nota: 0|5 representa 0,5 g

Diagrama de tallo y hoja del contenido de grasa saturada en 3 onzas de pescado.

0	0	1	1	2	2	3	4
0*	6						
1	1	1	2	4			
1*							
2							
2*	7						

Nota: 0|1 representa 0,1 g

Diagrama de tallo y hoja del contenido de colesterol en 3 onzas de pescado.

2	2			
3	5			
4	0	2	3	5
5	7	7		
6	0	1		
7	2			
8				
9	3			
10				
11				
12				
13				
14				
15				
16	6			

Nota: 2|2 representa 22 mg

Diagrama de tallo y hoja del contenido de calorías en 3 onzas de pescado.

```

7 | 5
8 | 2 3 4
9 |
10 | 9
11 | 7 9
12 | 6 7
13 | 2
14 |
15 | 5 7
16 |
17 |
18 |
19 | 6
  
```

Nota: 7|5 representa 75 calorías

Manos a la Obra N° 5

TABLA DE FRECUENCIAS DE LA ALTURA DE LAS MUJERES

Clase	Punto Medio	Frecuencia Absoluta	Frecuencia Relativa	Frecuencia Relativa Porcentual (%)	Frecuencia Absoluta Acumulada	Frecuencia Relativa Acumulada
[1,45-1,50)	1,475	1	0,02	2	1	0,02
[1,50-1,55)	1,525	6	0,12	12	7	0,14
[1,55-1,60)	1,575	18	0,36	36	25	0,50
[1,60-1,65)	1,625	18	0,36	36	43	0,86
[1,65-1,70)	1,675	5	0,10	10	48	0,96
[1,70-1,75)	1,725	2	0,04	4	50	1,00
Total	-	50	1,00	100	-	-

TABLA DE FRECUENCIAS DEL PESO DE LAS MUJERES

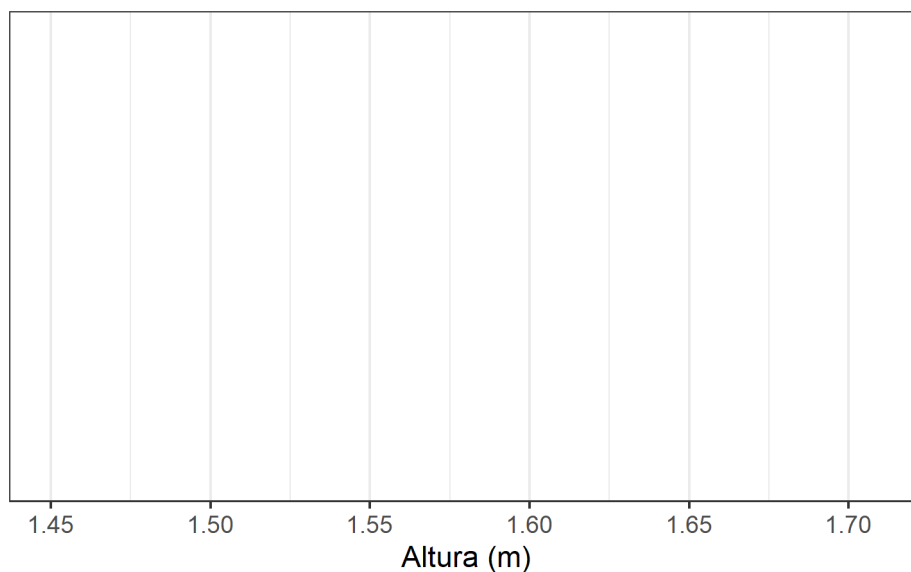
Clase	Punto Medio	Frecuencia Absoluta	Frecuencia Relativa	Frecuencia Relativa Porcentual (%)	Frecuencia Absoluta Acumulada	Frecuencia Relativa Acumulada
(35-40]	37,5	1	0,02	2	1	0,02
(40-45]	42,5	7	0,14	14	8	0,16
(45-50]	47,5	4	0,08	8	12	0,24
(50-55]	52,5	9	0,18	18	21	0,42
(55-60]	57,5	8	0,16	16	29	0,58
(60-65]	62,5	11	0,22	22	40	0,80
(65-70]	67,5	7	0,14	14	47	0,94
(70-75]	72,5	1	0,02	2	48	0,96
(75-80]	77,5	0	0,00	0	48	0,96
(80-85]	82,5	1	0,02	2	49	0,98
(85-90]	87,5	1	0,02	2	50	1,00
Total		50	1,00	100	-	-

Manos a la Obra N° 6HISTOGRAMA DE LA ALTURA DE LAS MUJERES

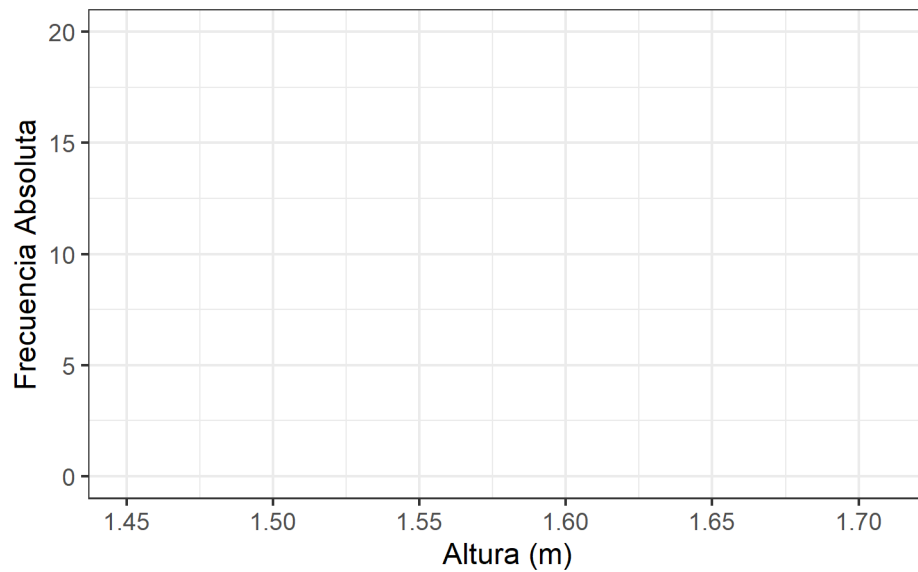
Usamos la tabla de frecuencias de las mujeres para elaborar el histograma de la altura de 50 mujeres.

Clase	Punto Medio	Frecuencia Absoluta	Frecuencia Relativa	Frecuencia Relativa Porcentual (%)	Frecuencia Absoluta Acumulada	Frecuencia Relativa Acumulada
[1,45-1,50)	1,475	1	0,02	2	1	0,02
[1,50-1,55)	1,525	6	0,12	12	7	0,14
[1,55-1,60)	1,575	18	0,36	36	25	0,50
[1,60-1,65)	1,625	18	0,36	36	43	0,86
[1,65-1,70)	1,675	5	0,10	10	48	0,96
[1,70-1,75)	1,725	2	0,04	4	50	1,00
Total	-	50	1,00	100	-	-

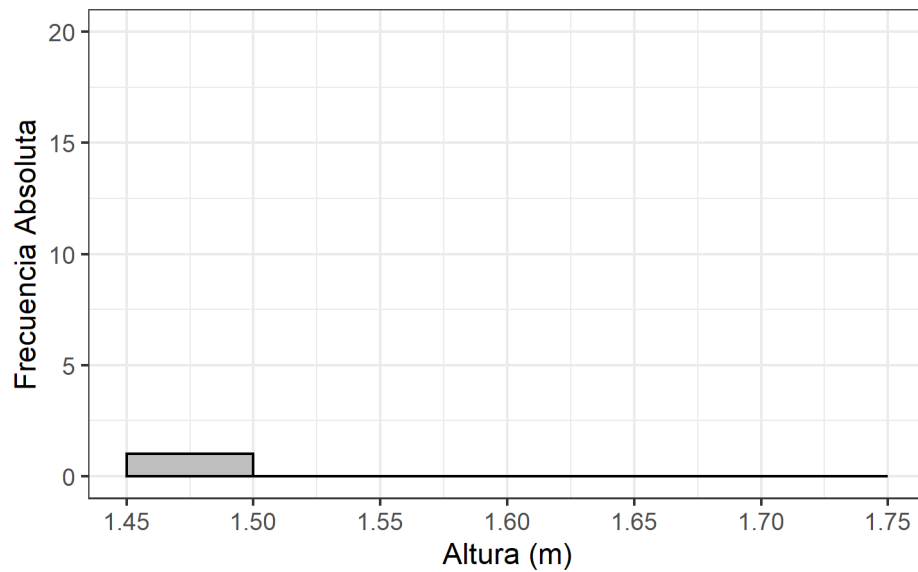
Empecemos con el histograma de las frecuencias absolutas. En el eje “x” ubicamos la variable en estudio:



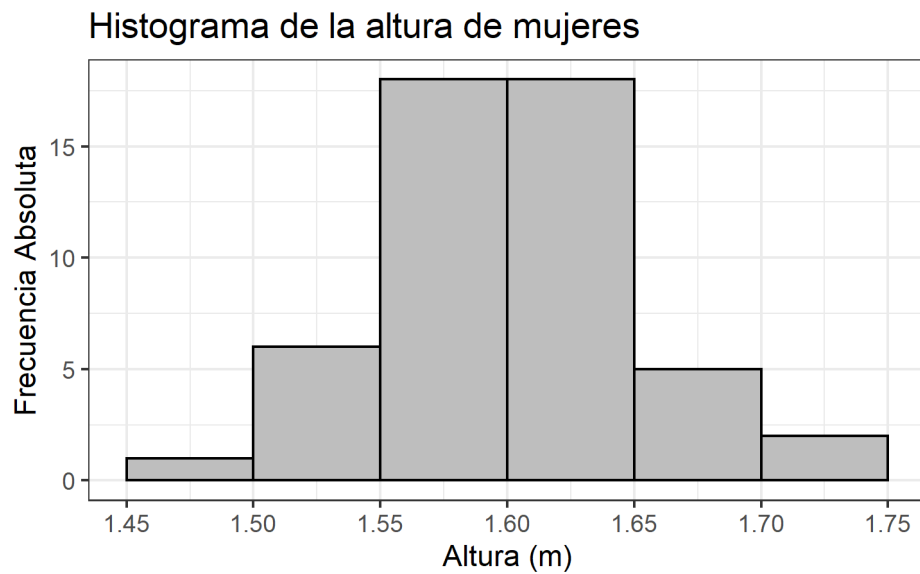
En el eje “y” vamos a representar la frecuencia absoluta. Como valor máximo del eje, tomamos un valor cercano a la máxima frecuencia absoluta observada, 18.



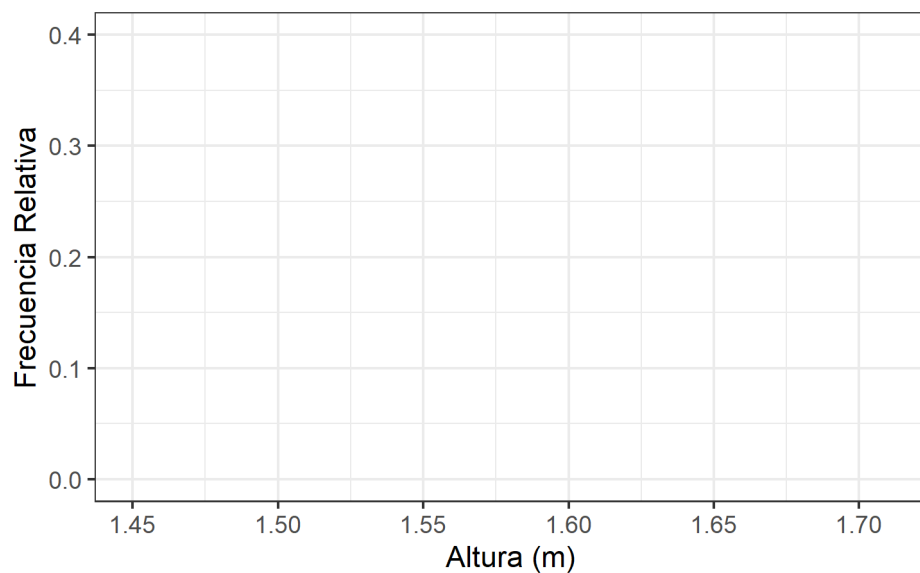
Ahora nos faltan las barras. La primera barra es un rectángulo cuya base empieza en el intervalo 1,45-1.50, y que tiene una altura que va hasta la frecuencia absoluta de 1:



Así vamos construyendo cada una de las barras para cada uno de los intervalos, y obtenemos el histograma. ¡Recordemos siempre agregar el título!

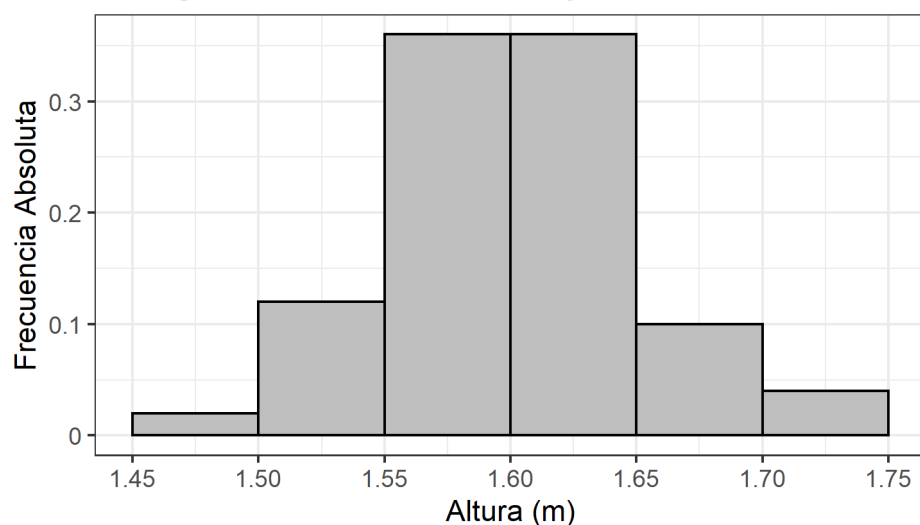


También podríamos graficar la frecuencia relativa. En ese caso, en el eje “y” representamos las frecuencias relativas:



El histograma resultante es:

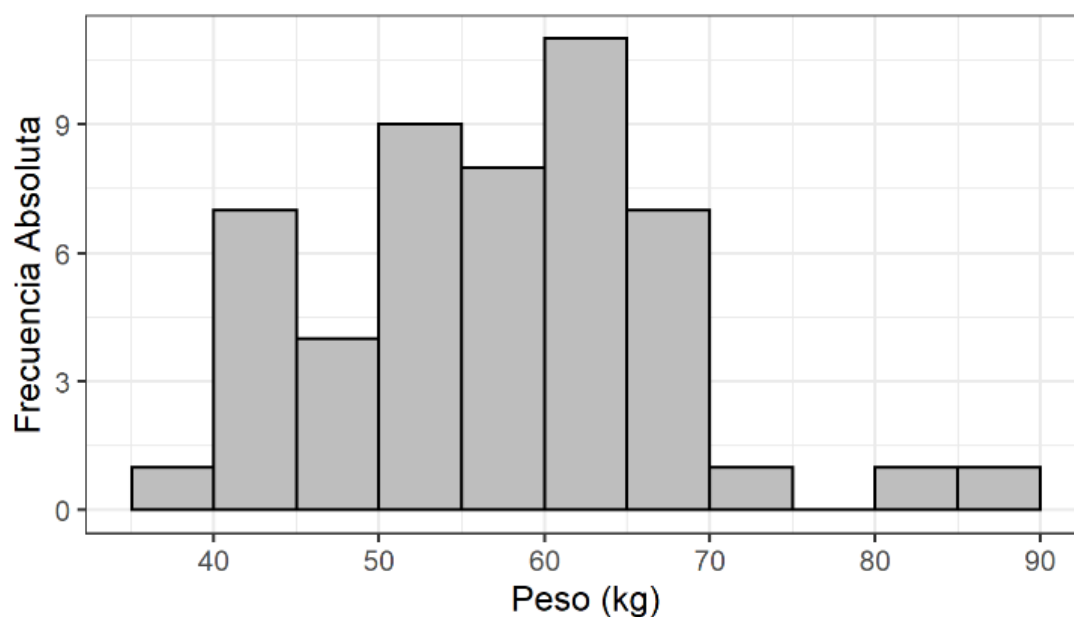
Histograma de la altura de mujeres



Las mujeres estudiadas tienen alturas entre 1,45 m y 1,75 m. La distribución de la altura es aproximadamente simétrica, y la mayoría de las mujeres miden entre 1,55 y 1,65 m.

HISTOGRAMA DEL PESO DE LAS MUJERES

Histograma del peso de las mujeres



Las mujeres analizadas tienen pesos entre 35 kg y 90 kg. La distribución del peso es ligeramente asimétrica a la derecha, sólo hay 3 mujeres con pesos mayores a 70 kg. La mayoría de las mujeres que participaron del estudio pesan entre 40 kg y 70 kg.