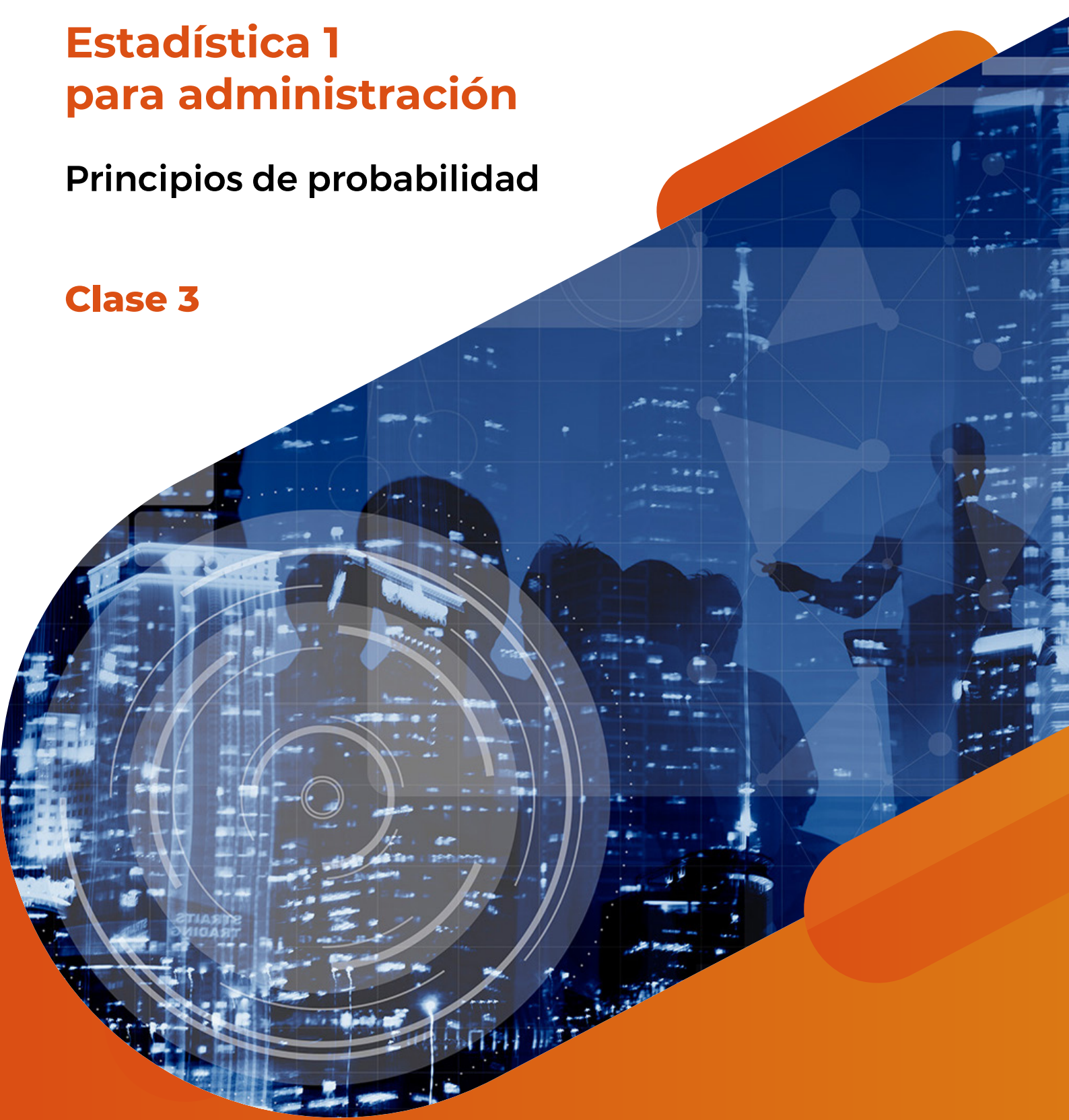


Estadística 1 para administración

Principios de probabilidad

Clase 3



INTRODUCCIÓN DE LA CLASE 3

RDA 1: Analizar información de un conjunto de datos, mediante tablas, gráficos y medidas descriptivas.

En esta clase iniciaremos con el análisis de la teoría de probabilidades. Continuamente los ejecutivos de una empresa se enfrentan a la toma de decisiones de negocios y en general sin importar la naturaleza del negocio es claro que dichas decisiones se basarán en la probabilidad de que determinado resultado positivo o negativo se produzca. Por otro lado, como habíamos visto en clases anteriores, este tipo de análisis estará siempre sujeto a un cierto nivel de incertidumbre.

Esta clase inicia con mostrar las distintas formas de medir la probabilidad de que un evento se presente, éstas varían desde situaciones simples a complejas, desde situaciones que involucran variables numéricas a categóricas y situaciones que involucran una muestra o población.

Principios de probabilidad

Iniciaremos con algunas definiciones

Precisión

De acuerdo a www.usc.gal Es la proximidad entre las indicaciones o los valores medidos, obtenidos en mediciones repetidas de un mismo objeto, bajo condiciones especificadas. La precisión se puede expresar numéricamente mediante medidas de dispersión tales como desviación típica, variancia o el coeficiente de variación bajo las condiciones especificadas. La precisión, se utiliza para definir a la repetibilidad de medida.

Ejemplo quiero medir la temperatura de un vaso de agua, realizo 20 mediciones, no todas sacan el mismo valor, hay décimas de grado de diferencia debido a la precisión del instrumento de medida.

Exactitud

En primer lugar, vale la pena puntualizar que la exactitud, (a diferencia de la precisión), no se expresa en forma numérica, sino que más bien es la diferencia del valor obtenido con algo que se considera verdadero. Por ejemplo, mido la temperatura del aire con un termómetro y la comparo con la temperatura publicada en una página web que se considera “correcta”. Por lo tanto, si lo comparo con otra página web, mi medición podría no ser tan correcta como pensaba.

Incertidumbre

De acuerdo a www.usc.gal La Incertidumbre es el parámetro asociado con el resultado de una medición, que caracteriza la dispersión de los valores que podrían ser razonablemente atribuidos al valor a medir. El valor de incertidumbre incluye componentes procedentes de efectos sistemáticos en las mediciones...

!Definición un tanto complicada! Más simple, cuando se desea medir un parámetro de la población, lo haremos mediante muestreo como ya se ha visto. El nivel de variabilidad entre una muestra y otra es la incertidumbre. Por ejemplo: Tengo un dataset con personas y datos relacionados a la persona. Realizo un muestreo y obtengo media de la edad de la muestra, digamos 35 años, luego realizo otro muestreo y obtengo 45 años, luego realizo otra medida y obtengo 40 años, puedo concluir que incertidumbre esta alrededor de 10 años.

Espacio Muestral

El espacio muestral asociado a un experimento dado es el conjunto de todos los posibles resultados que pueden surgir de su realización. Generalmente es designado con la letra E y colocamos sus posibles resultados entre llaves. Así el lanzamiento de 2 dados a la vez se denota:

$$E=\{(1,1),(1,2), (1,3)\dots\}$$

La probabilidad de que uno de los resultados del espacio muestral se presente es 1 y lo denotamos de la siguiente forma: $\sum P(E_i)=1$

Sucesos

Un suceso es toda proposición lógica que, una vez realizado el experimento aleatorio, se puede decirse se verifica o no, en otras palabras, podemos aseverar si la proposición fue verdadera o falsa

Probabilidad

De acuerdo a Wasserman (2015). En el lenguaje formal de la incertidumbre, que es la base de la inferencia, el problema básico que estudiamos en probabilidad es: dado un proceso generador de datos, cuáles son las propiedades del resultado. Más específicamente, que tan seguido podemos esperar obtener un valor en particular si repetimos una medición muchas veces y se denota como $P(X)$

La probabilidad se define como un valor numérico que estará entre 0 y 1 que representa cual es la posibilidad de que un evento se presente y se escribe de la siguiente forma: $P(x) = 0 \leq x \leq 1$, donde x representa el evento. Por ejemplo, la probabilidad de pasar este curso sin estudiar sería:

$$P(\text{sin estudiar})=0.$$

En general podemos establecer la siguiente fórmula para calcular una probabilidad:

Figura 1

Fórmula Base de la probabilidad Prado A. (2025)

Adaptación de la fórmula para la resolución de probabilidades de Webster (2000) capítulo 4

$$P(E) = \frac{\text{número de resultados favorables}}{\text{número total de posibles resultados}} = \frac{n(E)}{n(S)}$$

Existen básicamente 3 formas de cálculo de probabilidades:

- ✓ Modelos de frecuencia relativa o a-posteriori, basados en información histórica
- ✓ Modelo clásico o a priori, basado en posibles resultados
- ✓ Modelos subjetivos, cuando asignamos una probabilidad a algo que no sucedido

FRECUENCIA RELATIVA

Este modelo utiliza los datos de las observaciones que se han registrado, en base a estos calcula la frecuencia con la que se ha presentado dicho evento y en base a esta frecuencia se calcula la probabilidad.

Figura 2

Fórmula de frecuencia relativa Prado A. (2025)

Adaptación de la fórmula para la resolución de probabilidades de Webster (2000) capítulo 4

$$P(E) = \frac{\text{Número de veces que ha ocurrido el evento}}{\text{Número total de observaciones}}$$

Este modelo es también llamado a posteriori, dado que la data ya ha sido recabada. Por ejemplo, un gerente está preocupado por la cantidad de partes defectuosas E que recibe de un proveedor por caja y ha compilado la siguiente tabla

Tabla 1

Probabilidad a posteriori

Creación del autor Alfonso Prado

Resultado E	Número de cajas	P(E) Probabilidad relativa
0	40	0.4
1	27	0.27
2	21	0.21
3	12	0.12
Total	100	1

Este modelo puede presentar algunos problemas, por ejemplo, de la fórmula podemos deducir que, si las observaciones no incluyen una o más posibilidades (por ejemplo 4 defectos) será

imposible su cálculo de probabilidad y, por otro lado, si se cuenta con pocas observaciones sus resultados pueden ser engañosos.

PROBABILIDAD ADITIVA Y MULTIPLICATIVA

Es común que se pregunte por la probabilidad combinada de varias observaciones, para lo cual distinguiremos dos tipos de combinaciones:

Aditiva : Las probabilidades aditivas resultan de la suma de dos probabilidades, es decir, al investigador le da igual que cualquiera de los dos resultados ocurran. La probabilidad aditiva de A o B denotado como $P(A \cup B)$ y su fórmula es: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Por ejemplo: La probabilidad de sacar un as o una de las 13 cartas de corazones sería $P(\text{as} \cup \text{corazones}) = (4/52) + (13/52) - (1/52) = 16/52$

De acuerdo a Webster (2017) Los eventos A y B no necesariamente son excluyentes, es decir si ambos pueden ocurrir al mismo tiempo caeremos en doble conteo, en cuyo caso restaremos las posibilidades de este doble conteo ya que existe una carta que es as de corazones. Pero si los eventos son excluyentes entonces la fórmula quedaría así:

$$P(A \cup B) = P(A) + P(B)$$

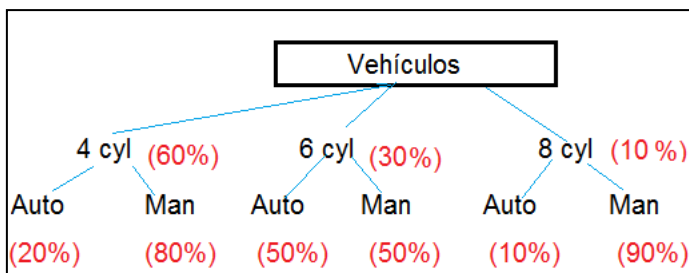
Multiplicativa: Denotada por: $P(A \cap B) = P(A) * P(B)$

Por ejemplo, del dataset mtcars, se desea obtener la probabilidad de obtener aleatoriamente un vehículo de 4 cilindros y que sea automático. Analizamos la distribución del dataset

Figura 3

Distribución de probabilidades de variables cyl y am en mtcars

Creación del autor Alfonso Prado



Entonces $P(4\text{cyl}) * P(\text{automático}) = 0.6 * 0.2 = 12\%$

Probabilidad condicional

Con frecuencia se desea determinar la probabilidad de algún evento (A), dado que antes otro evento ya haya ocurrido (B) y se denota como $P(A|B)$. Veamos algunos ejemplos:

Se sabe que una pareja tiene 2 hijos y que al menos 1 de ellos es niño. Cuál es la probabilidad de que los 2 sean varones? , En el lanzamiento de dados, ¿cuál es la probabilidad de sacar un 2, sabiendo de antemano de que la anterior fue un número par?

Figura 4

Fórmula de la Probabilidad condicional Prado A. (2025)

Adaptación de la fórmula para la resolución de la probabilidad condicional de Webster (2000) capítulo 4

Creación del autor Alfonso Prado

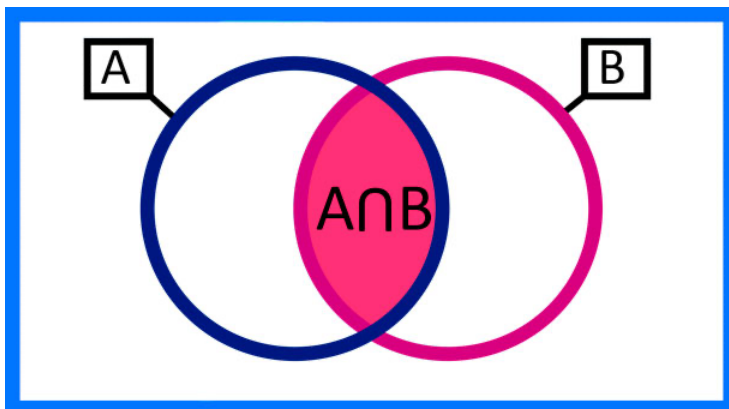
$$P(A | B) = \frac{P(A \cap B)}{P(B)} = P(A) * \frac{P(B|A)}{P(B)}$$

Como podemos ver el denominador es la intersección de 2 conjuntos. Aquí podemos tener dos posibilidades: Si los sucesos A y B son incompatibles (excluyentes) entonces la solución es fácil ya que A nunca se realizará. Caso contrario si $P(A \cap B)$ no es un conjunto vacío es posible que el evento A se pueda presentar. Sin embargo, el espacio muestral E se restringirá al evento B.

Figura 5

Diagrama de Venn de la probabilidad condicional

Creación del autor Alfonso Prado



La probabilidad condicionada permite calcular la probabilidad de la intersección de dos sucesos, es decir, la probabilidad de que se den ambos sucesos A y B, de esta forma, la

probabilidad de que tanto A como B ocurran es igual a la probabilidad de que A ocurra dado que B haya ocurrido multiplicado por la probabilidad de que B ocurra.

Por ejemplo, calcular cual es la probabilidad de obtener una J, dado que sabemos que es una figura F, denotado como $P(J|F)$

Aplicando la fórmula anterior tendríamos

$$P(J) = 4/52 \text{ dado que existen 4 Js}$$

$$P(F|J) = 1 \text{ porque todas las J son figuras}$$

$$P(F) = 12/52 \text{ porque existen 12 figuras.}$$

$$\text{Por lo tanto } P(J|F) = (4/52) * 1 / (12/52) = 4/12$$

Este es el concepto de probabilidad condicional del evento A dado que se conoce que el evento B ya ocurrió.

Teorema de Bayes

Supongamos que tenemos un conjunto completo de sucesos A_i , $i = 1, \dots, n$ y un suceso B cualquiera del espacio muestral. A veces es necesario conocer la probabilidad de uno de los sucesos A_j condicionada a que haya ocurrido B.

Veamos un ejemplo concreto:

Una fábrica tiene 2 equipos que producen la misma mercadería, la maquina 1 es más nueva y por lo tanto trabaja más rápido y produce el 60% de la producción, la maquina 2 es más antigua y produce el 40% restante. La máquina 1 además tiene apenas un 2% de producto descartado, mientras que máquina 2 tiene un 4% de producto descartado. De esta descripción podemos concluir lo siguiente tabla

Tabla 2
Cálculo de probabilidad del ejemplo
Creación del autor Alfonso Prado

Probabilidad	Fórmula
Máquina 1 Producto OK	$P(M1 \cap OK) = P(A) * P(OK A) = 0.6 * 0.98 = 0.588$
Maquina 1 Producto Des	$P(M1 \cap Des) = P(A) * P(Des A) = 0.6 * 0.02 = 0.012$
Maquina 2 producto Ok	$P(M2 \cap OK) = P(M2) * P(OK M2) = 0.4 * 0.96 = 0.384$
Máquina 2 Producto Des	$P(M2 \cap Des) = P(M2) * P(Des M2) = 0.4 * 0.04 = 0.016$

Entonces conociendo en que máquina fue producido un ítem podemos obtener su probabilidad que éste sea OK o Descartado. Sin embargo, es posible que deseamos hacer el análisis inverso: Dado un ítem que sabemos ha sido descartado, cual es la probabilidad de

que haya sido producido en la maquina 1. Para lo cual aplicamos la fórmula de probabilidad condicional figura 5.

Que en nuestro ejemplo sería:

$$P(M1 | Des) = \frac{P(M1 \cap Des)}{P(Des)} = P(M1) * \frac{P(Des | M1)}{P(Des)}$$

Pero el problema es que P(Des) no se conoce. Para resolver el problema partimos de la formula de probabilidad condicional (en la segunda forma) sustituyendo el evento B por D (descarte)

$$P(A|D) = P(A) \times P(D|A) / (P(D))$$

La probabilidad de descarte podría ser entendida por la regla de adición

$$P(D) = P(A \cap D) + P(B \cap D) = P(A) * P(D|A) + P(B) * P(D|B)$$

Cuando se hace la sustitución en el denominador de la fórmula de la probabilidad condicional anterior para P(D), el teorema de Bayes nos indicara:

$$P(D|A) = P(A \cap D) / (P(A \cap D) + P(B \cap D))$$

O lo que es igual a

$$P(A|D) = P(A) * P(D|A) / (P(A) * P(D|A) + P(B) * P(D|B))$$

Que en nuestro ejemplo sería

$$P(Des) = P(M1 \cap Des) / (P(M1 \cap Des) + P(M2 \cap Des))$$

$$P(Des) = 0.012 / (0.012 + 0.016) = 0.429$$

En general el teorema de Bayes puede ser generalizado para trabajar con k sucesos a la vez. Si se dispone de una serie de sucesos definidos como $\{A_i\}_{i=1...k}$ es decir un sistema completo de k sucesos, el teorema se generaliza de acuerdo a la siguiente figura.

Figura 6

Fórmula el teorema de Bayes generalizado Prado A. (2025)

Adaptación de las fórmulas para la resolución del Teorema según Soto (s.f.)

$$P(A_j / B) = \frac{P(A_j \cap B)}{P(B)} = \frac{P(A_j) \cdot P(B / A_j)}{\sum_{i=1}^{i=k} P(A_i) \cdot P(B / A_i)}$$

Teorema de Bayes

Contiene casos de uso y ejemplo

Representaciones básicas de variables cualitativas

Como habíamos mencionado anteriormente una variable cualitativa es aquella que no se expresa en forma numérica. Estas pueden ser nominales u ordinales, pero lo importante es que estas variables sirven como elementos de agrupación o categorización, por lo tanto, la representación de este tipo de variable se realiza en términos de la frecuencia de ocurrencia de dicha variable.

Existen varias representaciones para variables cualitativas:

- ✓ Diagramas de pastel
- ✓ Diagramas de mosaico
 - Diagramas de árbol
- ✓ Diagramas de ejes paralelos

Diagramas de pastel

El objetivo de estos gráficos es entender como se descompone un total en partes individuales. El gráfico típico en este caso es el diagrama de pastel, aunque el mismo procedimiento se puede realizar en un rectángulo y el resultado es un gráfico de barras apiladas. Los gráficos de pastel funcionan bien cuando el dataset es pequeño y no se requiere realizar comparaciones periódicas (como series de tiempo). Si se requiere comparar pasteles las diferencias pueden ser difícil de notar. La comparación se vuelve un poco más clara cuando usamos barras apiladas, aunque (otra vez) si el conjunto es grande la participación de cada proporción en el tiempo sigue siendo difícil de seguir. Es importante mencionar que estos diagramas no están hecho para mostrar valores (frecuencias) exactas

Diagramas de Mosaico

No es raro que queramos profundizar más y desglosar un conjunto de datos por múltiples variables categóricas a la vez. Por ejemplo, considere la data del Titanic. El dataset está compuesto por múltiples variables categóricas por lo que utilizar un diagrama de pastel no sería lo más adecuado, en su lugar haríamos un mosaico. En este tipo de gráfico, el área de cada una de las secciones representa la frecuencia de ocurrencia de la variable categórica.

Tabla 3

Pasajeros de Titanic

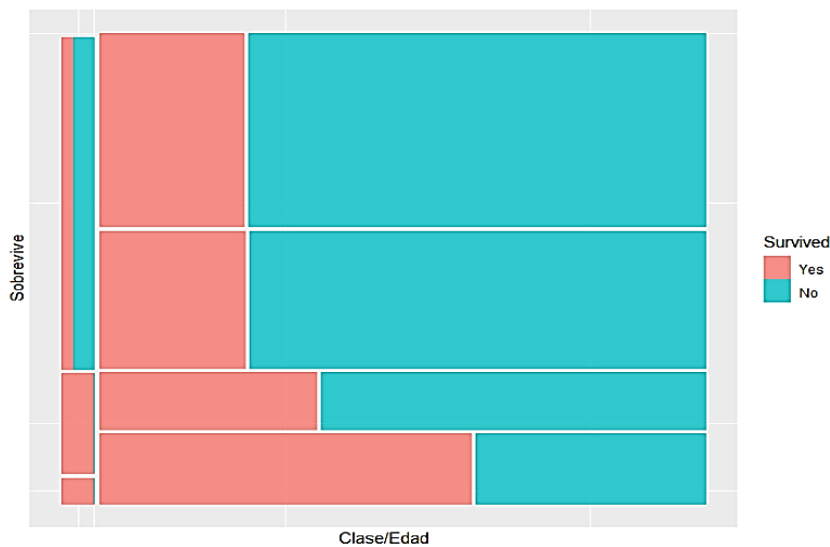
Creación del autor Alfonso Prado

Gender	Survived	1st Class	2nd Class	3rd Class	Crew
Male	No	118	154	422	670
	Yes	62	25	88	192
Female	No	4	13	106	3
	Yes	141	93	90	20

En estos casos podemos tener dos posibilidades: Que un elemento pueda estar contenido en dos o más categorías o que cada elemento pertenezca exclusivamente a un categoría y que las categorías se anidan unas dentro de otras. Este último es conocido como un diagrama de árbol (treemap)

Figura 7

Diagrama de mosaico con múltiples variables categóricas del dataset Titanic
Creación del autor Alfonso Prado

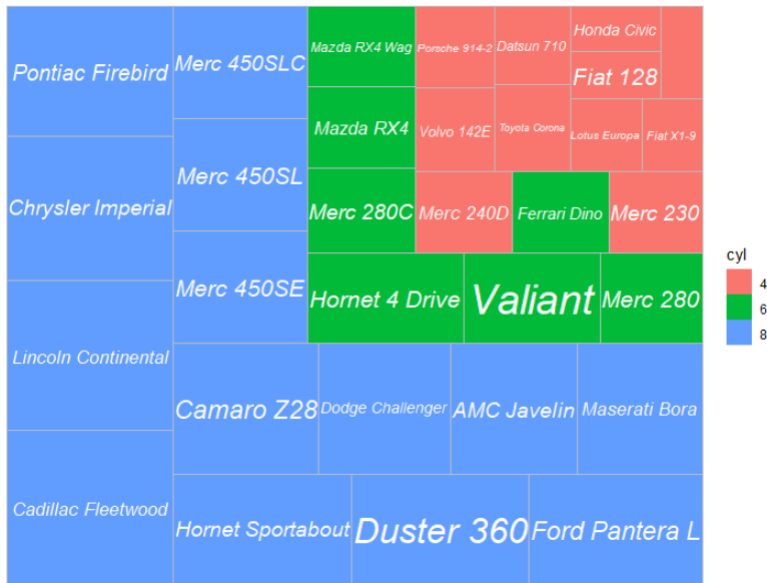


Los mapas de árbol muestran datos jerárquicos (estructurados en árbol) como un conjunto de rectángulos anidados. A cada rama del árbol se le da un rectángulo, que luego se coloca en mosaico con rectángulos más pequeños que representan subramas. Pero note que no hay posibilidad que un factor esté en más de un lugar.

Considere la data de mtcars, la variable categórica es cyl que representa la cantidad de cilindros que tiene el motor (4,6 u 8), por lo que no hay posibilidad de un vehículo se halle en dos categorías a la vez.

Figura 8

Diagrama de árbol del dataset mtcars
Creación del autor Alfonso Prado



Diagramas de ejes paralelos

En los gráficos de conjuntos paralelos, a cada variable categórica se le asignará una posición en el eje X. El tamaño de la intersección de categorías de variables vecinas se muestra luego como diagonales gruesas, escaladas por la suma de elementos compartidos entre las dos categorías.

Diagramas de ejes paralelos
Describe su uso y ejemplos

REFERENCIAS BIBLIOGRAFICAS

Conceptos de estadística , (<https://www.usc.gal/genp/docencia/ConceptosDeEstadistica.pdf>)
 Wasserman, L. (2015). All of Statistics. Springer Texts
 Gamero Burón, C. (2017). Estadística I: elementos de estadística descriptiva y de teoría de la probabilidad. Universidad de Málaga.
 Soto, G. (s.f.) El teorema de Bayes , Revista de la Universidad Nacional de Córdoba 26-3



La excelencia no se improvisa

síguenos

