



UNIVERSIDAD DE ATACAMA
FACULTAD DE INGENIERÍA / DEPARTAMENTO DE MATEMÁTICA

ESTADÍSTICA Y PROBABILIDADES

GUÍA N°1: Estadística y Probabilidades

Profesor: Hugo S. Salinas.

Primer Semestre 2011

-
1. Señalar el tipo de variable que corresponde: Cuantitativa (continua, discreta) o Categórica (Ordinales, Nominales):

Edad:

Género:

Número de hijos:

Altura:

Número de ataques de asma por semana:

De acuerdo, neutral, en desacuerdo:

Presión arterial:

Bien, igual, peor:

Vivo o muerto:

2. Un Ingeniero Civil visita 25 villas en una ciudad y en cada una registró el número de casas que han sufrido daños ocasionados por un terremoto, de lo cual resultaron los datos:

15 20 25 15 18 16 17 18 20 18 18 18 19 16 17 19 16 17 17 17 19 18 19 18 15

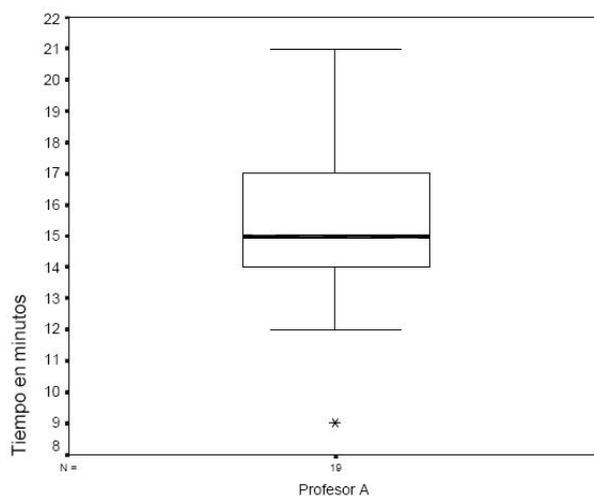
- ¿Qué tipo de datos son estos?
 - Construir una tabla de distribución de frecuencias adecuadas a este conjunto de datos.
 - ¿Cuántas villas tienen a lo más 20 casas que han sufrido daños?
 - ¿Qué proporción de villas tienen por lo menos 17 casas que han sufrido daños?
 - ¿Qué proporción de villas tienen 18 casas que han sufrido daños?
 - ¿Qué proporción y que porcentaje de villas tienen 18 o menos casas que han sufrido daños?
 - Calcular e interpretar la media aritmética de los datos.
 - Obtener e interpretar la mediana de los datos.
 - Construir un gráfico adecuado y comentar.
3. Los datos siguientes representan una muestra de 80 determinaciones de la emisión diaria de óxidos de azufre de una planta industrial (en toneladas), ordenados de menor a mayor:

6.2 7.7 8.3 9.4 10.5 ... 29.6 31.8 35.0 39.8

A continuación se resumen algunos estadísticos calculados sobre esta muestra:

Máximo = 39.8
 Mínimo = 6.2
 Cuartil 1 = 14.95
 Cuartil 2 = 19.05
 Cuartil 3 = 22.95
 Promedio = 21.5

- a) Interpretar los valores del cuartil 2 y del cuartil 3
 - b) Representar en un boxplot, el resumen de los 5 números e interprete.
 - c) Determinar si existen datos anómalos (outliers) en la muestra.
4. Dos profesores (*A* y *B*) están interesados en estudiar los hábitos de sueño de los estudiantes en sus clases. Ambos profesores registran el tiempo (en minutos) que demoran en quedarse dormidos sus alumnos desde que empieza la clase. El gráfico muestra los tiempos que demoran en quedarse dormidos los alumnos del profesor *A*.



- a) ¿Cuál es el valor aproximado de las medidas de dispersión del tiempo del Profesor *A*?
- b) ¿Qué porcentaje de alumnos se queda dormido antes de los 14 minutos con el Profesor *A*? Justificar.
- c) Los datos del Profesor *B* son los siguientes:

10.5	11.3	11.9	12	12.3	12.3	12.5	12.7	13.4	13.7
13.8	14.2	14.8	15.1	15.3	16.7	16.8	18.8	20.8	

Construir un diagrama de cajas correspondiente a los tiempos en que se quedan dormidos los alumnos en la clase del Profesor *B*.

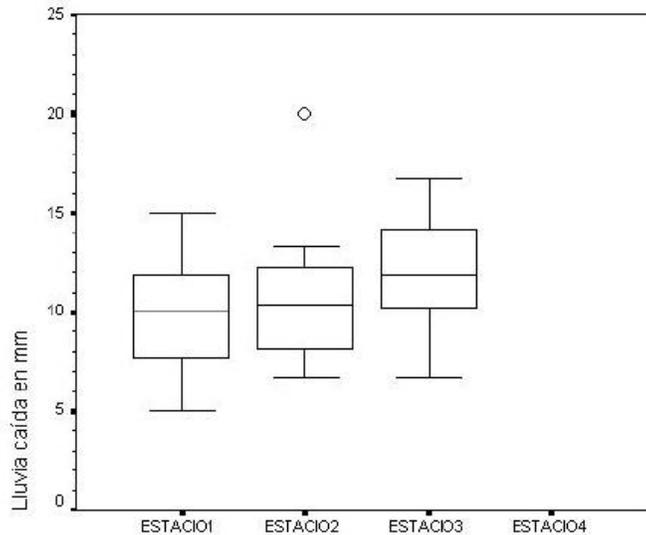
5. En el ANEXO hay una base de datos con información de 36 alcaldes de diversas comunas de Chile, para cada uno se presentan las siguientes variables:

Pacto político del alcalde.
 Años que lleva en el cargo el alcalde.
 Número de veces electo.
 Nivel de estudios alcanzado.
 Edad del alcalde en años.

- a) Clasificar cada una de las variables.
- b) Realizar una tabla de frecuencia univariada (por cada variable).
- c) Interpretar al menos una fila por cada tabla.

d) Realizar un análisis estadístico descriptivo (con lo visto en clases)

6. Se tienen datos de la cantidad mensual de lluvia caída en 3 estaciones meteorológicas de una ciudad para los 12 meses de 1999.



- ¿Cuál de las 3 estaciones tiene el mayor rango de lluvia caída en los 12 meses?. Justificar.
- En general, ¿Qué medidas de dispersión puedes calcular en un gráfico de caja?
- ¿Cuál es el valor aproximado de las medidas de dispersión en la estación 2?
- ¿Aproximadamente cuántos meses llovió menos de 8 mm en la estación 1. Justificar.
- En vista de que la caja de la estación 1 es simétrica, ¿Podemos decir que la distribución de los valores de lluvia caída en la estación 1 será simétrica?. Justificar.
- Estos son los datos de la lluvia caída en la estación 4:

3 5 9 9 9 10 11 11 11 11 12 15

- Calcular las medidas de tendencia central y de dispersión a estos datos.
- Dibujar en el gráfico la caja correspondiente a la estación 4.

7. Un corrector de textos contabiliza el número de erratas que encuentra en cada página. Después de pasar este corrector por un texto de 50 páginas, se obtiene el siguiente número de erratas por página:

2 3 5 0 1 4 0 6 2 1 1 0 2 4 5 3 1 2 3 2 3 1 2 4 4
2 5 4 1 3 2 6 8 2 0 1 0 2 3 1 5 10 2 1 3 6 2 0 1 3

- A partir del enunciado del problema, introducir una variable estadística y decidir de qué tipo es.
- Construir la tabla de frecuencias correspondiente.
- ¿Cómo sería el diagrama de barras de las frecuencias absolutas? Solamente mirando este gráfico, ¿podrás decir cuánto vale la moda?
- ¿Qué porcentaje de páginas, respecto del total de las que se han corregido, tienen 2 erratas?
- ¿Qué porcentaje de páginas respecto del total tienen menos de 6 erratas? ¿Y 6 erratas o más?
- ¿Qué porcentaje de páginas respecto del total tienen como mínimo 5 erratas?
- Calcular los siguientes estadísticos descriptivos: media aritmética, mediana, moda, varianza y cuartiles.

- h) Si comparas la media aritmética y la mediana, ¿qué puedes decir de la simetría de la distribución? ¿Entre qué valores se encuentran el 50% central de los datos?
- i) Construir un diagrama de caja. ¿Hay algún valor anómalo? ¿Cuál es? Fijarse en la asimetría de la distribución: ¿presenta una cola hacia los valores grandes o hacia los valores pequeños?

8. Los datos siguientes expresan el número de días transcurridos hasta la primera falla en cierto tipo de electrodoméstico:

534	875	435	654	432	984	321	765	453
765	564	982	873	567	871	658	564	399

- a) Calcular la media, desviación estándar, mediana y rango intercuartil de las observaciones.
- b) Encontrar la transformación lineal de la variable que representa el tiempo de duración en semanas.
- c) Obtener la media, desviación estándar, mediana y rango intercuartil de los datos transformados. ¿Qué relación tienen con los valores obtenidos usando los datos originales?
9. En un estudio sobre parásitos se consideró el número de garrapatas en el cuerpo de ratones. En un grupo de 44 ratones se obtuvieron los siguientes resultados:

Número de garrapatas:

0	1	1	1	0	2	3	4	0	5	0
0	0	0	1	0	0	0	2	3	2	1
1	2	0	2	0	4	1	1	0	0	2
1	0	0	1	0	0	1	1	0	0	6

Sexo:

M	H	H	H	H	M	M	M	M	H	H
H	H	H	M	M	H	M	H	H	M	M
M	H	H	H	M	M	H	H	H	M	H
H	M	M	H	H	M	M	H	M	M	H

- a) ¿De qué tipo son las variables número de garrapatas y sexo?
- b) Para la variable número de garrapatas calcular la media aritmética, la mediana, los cuartiles, la moda, el rango, la varianza, la desviación estándar, el **coeficiente de asimetría**¹ y el **coeficiente de curtosis**². ¿Qué información puedes sacar? ¿Para la variable sexo pueden calcularse también todas estas medidas?
- c) Construir la tabla de frecuencias para la variable número de garrapatas, de manera que los puntos medios de los intervalos de clase coincidan con los valores que toma la variable. ¿Cuántos ratones tienen exactamente 3 garrapatas? ¿Qué porcentaje de ratones tienen más de cuatro garrapatas? ¿Qué porcentaje de ratones tienen al menos una garrapata? Hacer la representación gráfica adecuada para este tipo de variable.
- d) Hacer un diagrama de caja para la variable número de garrapatas. ¿Hay valores anómalos?
- e) Construir la tabla de frecuencias para la variable sexo y hacer la representación gráfica correspondiente para este tipo de variable.
- f) Obtener la tabla de contingencias para la variable sexo en las filas y la variable número de garrapatas en las columnas. ¿Cuántos ratones macho tienen 4 garrapatas? ¿Qué porcentaje representan respecto del total de machos? ¿y respecto del total de ratones? ¿Cuántos ratones hembra tienen más de 5 garrapatas? ¿Qué porcentaje representan respecto del total de hembras? ¿Qué porcentaje representan respecto del total de ratones?

¹ $skewness = \frac{\sum_{i=1}^n (X_i - \bar{X})^3}{(n-1)S^3}$

² $kurtosis = \frac{\sum_{i=1}^n (X_i - \bar{X})^4}{(n-1)S^4}$

10. Describir el espacio muestral de las siguientes experiencias aleatorias:

- a) $E_1 = \{\text{Lanzamiento de un dado y anotamos el resultado}\}.$
- b) $E_2 = \{\text{Lanzamiento tres dados y sumamos las puntuaciones}\}.$
- c) $E_3 = \{\text{La duración de una ampolleta hasta que se funde}\}.$
- d) $E_4 = \{\text{La resistencia a rotura de unos tubos de aluminio}\}.$
- e) $E_5 = \{\text{Número de piezas defectuosas de un lote de 5000}\}.$
- f) $E_6 = \{\text{Lanzamiento de dos monedas}\}.$

11. Tres sucesos A , B y C verifican

$$\begin{aligned}P(A) &= P(B) = P(C) = 1/3 \\P(A \cap B) &= P(B \cap C) = P(A \cap C) = 1/9 \\P(A \cap B \cap C) &= 1/27\end{aligned}$$

Calcular

- a) La probabilidad de que ocurra exactamente uno de los tres eventos.
- b) La probabilidad de que ocurra al menos uno de los tres sucesos.

Solución: (a) $4/9$ (b) $19/27$

12. En la fabricación de cierto artículo se encuentra que se presenta un tipo de defecto con una probabilidad de 0.1, y también defectos de un segundo tipo con una probabilidad de 0.05. Sabiendo que ambos defectos son independientes, cuál es la probabilidad de que

- a) un artículo no tenga ambas clases de defectos.
- b) un artículo sea defectuoso.
- c) suponiendo que un artículo sea defectuoso, tenga sólo un tipo de defecto.

Solución: (a) 0.995 (b) 0.145 (c) 0.9655

13. Dados dos eventos A y B tales que $P(A) = 0.5$ y $P(A \cup B) = 0.7$. Calcular $P(B)$

- a) si A y B son independientes.
- b) si A y B son mutuamente excluyentes.
- c) si $P(A|B) = 0.5$.

Solución: (a) $2/5$ (b) 0.2 (c) $2/5$

14. 12. En una fábrica la máquina 1 produce piezas de buena calidad con una probabilidad igual a 0.8; la máquina 2 con probabilidad igual a 0.9. Se separa una pieza de cada máquina, calcular:

- a) la probabilidad de que ambas piezas sean defectuosas
- b) la probabilidad de que una sea defectuosa y la otra no

Solución: (a) 0.02 (b) 0.26

15. En un sistema de alarma, la probabilidad de que se produzca un peligro es 0.1. Si éste se produce, la probabilidad de que la alarma funcione es 0.95. La probabilidad de que funcione la alarma sin que haya peligro es 0.03. Hallar:

- a) la probabilidad de que habiendo funcionado la alarma, no haya habido peligro.
- b) la probabilidad de que haya un peligro y la alarma no funcione.

c) la probabilidad de que no habiendo funcionado la alarma, hay un peligro.

Solución: (a) 0.221 (b) 0.005 (c) 0.0057

16. Durante un año, las personas de una ciudad utilizan tres tipos de transportes: metro (M), micro bus (A) y auto particular (C). Las probabilidades de que durante el año hayan usado unos u otros transportes son las siguientes: $P(M) = 0.30$, $P(A) = 0.20$, $P(C) = 0.15$, $P(M \cap A) = 0.10$, $P(M \cap C) = 0.05$, $P(A \cap C) = 0.06$ y $P(M \cap A \cap C) = 0.01$. Calcular las probabilidades siguientes:

- a) que una persona tome al menos dos medios de transporte.
- b) que una persona viaje en metro y no en micro bus.
- c) que una persona viaje en metro o en auto particular y no en micro bus.
- d) que viaje en metro o en micro bus y en auto particular.
- e) que una persona vaya a pie.

Solución: (a) 0.19 (b) 0.20 (c) 0.25 (d) 0.10 (e) 0.55

17. La siguiente tabla muestra resultados parciales de una encuesta aplicada a 200 estudiantes de Ingeniería.

Fuma	Nivel de actividad física		
	Nula	Leve	Intensa
SI	60	40	20
NO	20	40	20

De acuerdo a lo anterior, determinar:

- a) Probabilidad de presentar un nivel leve de actividad física.
- b) Si es fumador, ¿cuál es la probabilidad que presente un nivel leve de actividad física?
- c) Probabilidad de fumar y tener un nivel leve de actividad física.
- d) Probabilidad que fume o presente un nivel leve de actividad física.

18. Tenemos cuatro urnas que contienen bolas de tres colores: blanco, negro y rojo, con las composiciones siguientes:

Urna	Blancas	Negras	Rojas
U	10	20	30
U_1	1	2	7
U_2	3	2	5
U_3	4	4	2

Se extrae una bola de una de las tres urnas U_1 , U_2 o U_3 , dependiendo la elección de la urna de que al extraer una bola de la urna U sea blanca, caso en el que se selecciona la urna U_1 , negra (urna U_2) o roja (urna U_3). Si se deseara apostar por un color, ¿cuál elegirías?.

Solución: rojo.

19. Una entidad bancaria, partiendo de la información que posee sobre el comportamiento de sus cuentacorrentistas, referente a los errores cometidos en los cheques extendidos por ellos, ha llegado a los siguientes resultados: de 850 clientes con fondos, ha habido 25 que pusieron algún error. El 98% de los clientes tienen fondos. De 50 cheques sin fondos, 45 tenían errores. Partiendo de esta información se desea hallar la probabilidad de que un cheque con errores resulte sin fondos.

Solución: 0.388

20. En un sistema de alarma, la probabilidad de que se produzca una situación de peligro es 0.10. Si esta tiene lugar, la probabilidad de que la alarma funcione es 0.95. La probabilidad de que funcione la alarma sin que haya situación de peligro es 0.03.

Sean $S := \{\text{se da una situación de peligro}\}$ y $F := \{\text{funciona la alarma}\}$. Determinar:

- La probabilidad de que habiendo funcionado la alarma no haya situación de peligro.
 - La probabilidad de que se de situación de peligro y la alarma no funcione.
 - La probabilidad de que no habiendo funcionado la alarma, haya peligro.
21. Una empresa minera busca nuevos yacimientos de cobre. Un yacimiento es explotable si contiene al menos cierta cantidad de cobre en la misma localización. Se ha realizado una excavación en una zona para determinar la presencia de cobre. Un estudio geológico ha estimado que la probabilidad de que un yacimiento sea explotable es 0.2. Si el yacimiento es explotable, la probabilidad de que se descubra cobre en la excavación es de 0.8 y si no, la probabilidad es 0.1.
- Calcular la probabilidad de que se descubra cobre en la excavación.
 - Dado que se descubre cobre, calcular la probabilidad de que el yacimiento sea explotable.
 - Dado que no se descubre cobre, calcular la probabilidad de que el yacimiento sea explotable.
22. Considerar una sucesión de días y sea el evento $R_i := \text{llueve el día } i$. Supongamos que

$$P(R_i|R_{i-1}) = \alpha \quad \text{y} \quad P(R'_i|R'_{i-1}) = \beta$$

Además, supongamos que sólo el clima de hoy es relevante para predecir el clima de mañana.

- Si la probabilidad de que llueva hoy es p . ¿Cuál es la probabilidad de que llueva mañana?
 - ¿Cuál es la probabilidad de que llueva pasado mañana?
 - ¿Cuál es la probabilidad que lluevan los n días después de hoy?
23. Supongamos que se tienen 3 posters, de los cuales uno tiene en ambas caras al candidato A ; el otro tiene en ambas caras al candidato B y el tercero tiene en una cara al candidato A y en la otra al candidato B . Para colgar el poster en una pared, se extrae uno de los tres al azar y se coloca, también al azar una cara al frente.
- Determinar un espacio muestral para este problema.
 - ¿Cuál es la probabilidad de que el poster colocado en la pared muestre al candidato A ?
 - Si el poster muestra al candidato A , ¿Cuál es la probabilidad de que la cara contra la pared sea la del candidato B ?
24. Los ciudadanos expresan sus intenciones de votos en encuestas electorales. Supongamos que en el pasado, 95% de los candidatos con destacada actuación política recibieron buenas evaluaciones de parte del electorado; el 60% de los candidatos con razonable actuación política recibieron buenas evaluaciones y el 10% de los candidatos con mala actuación política recibieron buenas evaluaciones. Además, el 40% de los candidatos tuvieron destacada actuación política, el 35% una actuación política razonable, y el 25% una mala actuación política.
- ¿Cuál es la probabilidad de que un candidato tenga una buena evaluación?
 - Si el candidato tiene una buena evaluación, ¿Cuál es la probabilidad de que tenga una destacada actuación política?

ANEXO

Comuna	Pacto político del alcalde	Años en el cargo	Nº de veces electo	Estudios alcanzado	Edad
Arica	Independiente	4	1	superior	42
Iquique	alianza por chile	7	2	técnico	70
Calama	Independiente	2	1	superior	47
Antofagasta	Independiente	2	1	superior	50
Copiapo	Independiente	2	1	técnico	52
La Serena	Concertación	9	3	superior	59
Coquimbo	Concertación	4	1	técnico	57
Valparaíso	alianza por chile	2	1	técnico	52
Viña del Mar	alianza por chile	6	2	enseñanza media	71
Rancagua	alianza por chile	2	1	enseñanza media	51
Curico	alianza por chile	2	1	superior	36
Talca	alianza por chile	2	1	técnico	50
Constitución	Independiente	2	1	técnico	61
Linares	alianza por chile	2	1	técnico	42
Parral	Concertación	2	1	enseñanza media	61
Chillan	alianza por chile	2	1	superior	67
Concepción	alianza por chile	1	0	superior	54
Talcahuano	Independiente	2	1	superior	54
Temuco	alianza por chile	2	1	técnico	49
Villarrica	alianza por chile	2	1	superior	56
Padre de las casas	alianza por chile	2	1	superior	41
Valdivia	alianza por chile	10	3	superior	59
Osorno	Concertación	2	1	técnico	56
Puerto Montt	Concertación	10	3	superior	64
Puerto Varas	alianza por chile	6	2	superior	58
Castro	Concertación	14	4	superior	58
Coyhaique	alianza por chile	2	1	técnico	39
Punta Arenas	Independiente	2	1	superior	64
Santiago	alianza por chile	2	1	superior	47
Ñuñoa	alianza por chile	14	4	superior	56
Pudahuel	Concertación	18	5	superior	58
La Pintana	Concertación	18	5	enseñanza media	50
Macul	Concertación	18	5	enseñanza media	72
Providencia	alianza por chile	14	4	superior	53
Las Condes	alianza por chile	10	3	superior	61
Puente Alto	alianza por chile	10	3	técnico	48