

Universidad de Costa Rica
Sede Occidente
Departamento de Ciencias Sociales
Curso: Estadística general 1
Prof. Carlos Conejo
III semestre 2008.

TEMA 1. CONCEPTOS ESTADISTICOS

1.1 Significado de la palabra estadística. (datos, disciplina)

- Presentar *la portada* del libro de Oscar Hernández (datos, gráfico, fórmula)
- Todos hemos escuchado y usado la palabra estadística, sin precisión.
- No tenemos capacidad de recordar todos los datos de X variable
- El término se utiliza en todas las disciplinas científicas.
- Ejemplo: las estadísticas del partido, las estadísticas de ventas, empleo, muertes.
- Estadísticas como serie de datos vs estadística como disciplina:
- Se interesa por analizar los fenómenos *regulares* de un conjunto de datos.
- Estadística: rama de la matemática aplicada, técnicas e instrumentos.
- Mentiras estadísticas

- Definición: *disciplina encargada del desarrollo y aplicación de la teoría y las técnicas apropiadas para la recolección, clasificación, presentación, análisis e interpretación de la información, que puede ser cuantitativa o cualitativa.*
 - Teoría: modelos matemáticos
 - Técnicas: instrumentos de análisis numérico

- Estadística descriptiva e inferencial (inductiva)
 - Estadística descriptiva: se encarga de describir las características principales de grupos de datos (recolección, ordenamiento y presentación de datos). Ejem. Estudiar el rendimiento estudiantil de la carrera de Administración de empresas en Sede Occidente en 2008.
 - Estadística inferencial: Hacer generalizaciones de los resultados de una muestra al conjunto de la población de interés. Ejm. En mismo estudio anterior pero para identificar el rendimiento académico a nivel de educación superior en Costa Rica (incluye una muestra de grupos de estudiantes de todas las universidades que imparten la carrera de Administración de Empresas en el país.

- Estadística: disciplina que ofrece los métodos más adecuados para el análisis de conjuntos de datos. Con ella podemos, por ejemplo:
 - Calcular promedio: que da una indicación general de x variable
 - Determinar la variabilidad de los datos: homogeneidad
 - Preparar cuadros y gráficos: facilita la interpretación de
 - Es posible analizar relación entre variables: regresión y correlación

- Predicciones: a partir del comportamiento actual de los datos
- Generalizar resultados.

1.2 Algunos conceptos básicos:

- **Unidad estadística, característica, observación:** el análisis estadístico se lleva a cabo con base en *observaciones* que corresponden a cierta *característica* que posee determinada *unidad estadística* o unidades de estudio).
 - **Unidad estadística:** unidad de interés en el campo de estudio (persona, animal, institución, cosa)
 - **Característica:** elemento de interés en el estudio (variables, atributos)
 - Característica cuantitativa: variable (discreta, continua)
 - Característica cualitativa: atributo (color en carro)
 - **Observación:** Valor de la característica para x Unidad estadística
- Ejemplo 1: si al finalizar este curso se quisiera hacer un análisis estadístico del rendimiento académico del grupo, entonces primero se deben definir con claridad esos elementos:
 - UE = cada uno de los estudiantes (Pedro)
 - C = rendimiento académico (nota final)
 - Ob. = 80.
- Ejm 2: Condición de los préstamos del Banco Popular de San Ramón, en 08.
 - UE = cada préstamo,
 - C = estado (de cumplimiento),
 - Ob. = al día, atrasado, cobro judicial.
- Ejm 3: Desempeño de los 5 vendedores del residencial La Laguna 2008.
 - UE = cada uno de los 5 vendedores (Pedro, Ana, José, Juan, Juana)
 - C = Nivel de ventas por mes, y total anual
 - Ob.= Juan, 2 casas en el mes de julio
- Ejm. 4: determinar el rendimiento por kilometraje de los autos Yaris:
 - UE = cada uno de los autos Yaris a considerar
 - C = kilometraje por galón de gasolina
 - Ob. 50 km/galón el carro uno.
- Ejm. 5: producción de leche de las 20 vacas de la Finca X en diciembre:
 - UE = cada una de las 20 vacas
 - C = producción de leche en litros
 - Ob. = 4 libras (vaca 1, día 2)
- Ejem. 6: salario de los funcionarios de Sede Occidente:
 - UE = cada uno de los funcionarios
 - C = Salario mensual
 - Ob. 200,000

- Es importante definir con precisión estos aspectos, es especial la característica (variable, porque su valor varía de una UE a otra), en este caso, si se incluye el salario bruto o neto, si incluye horas extras, en cuál periodo, interinos o propietarios, tiempo completo o también incluye a los funcionarios que están contratados a tiempo parcial.
- Generalmente en un estudio se analizan varias características a la vez. Ejm del Director de colegio que quiere estudiar los elementos ligados al rendimiento académico de sus estudiantes, puede tener algo así, para un estudiante:

Unidad de estudio (est.)	Característica	Observación
Jorge Sánchez Vargas	Nota promedio, 1er trim.	85
	Sexo	M
	Ocupación del padre	Profesor
	Escolaridad del padre	Bachiller
	Ingreso mensual familiar	500 mil
	Reciben periódico en casa	Si
	Coficiente de inteligencia	102
	Test de conocimientos	70

Luego se hacen relaciones entre las observaciones de las diferentes características.

Población: conjunto de miembros que comparten una característica particular, en un estudio determinado.

- Ejem. Los estudiantes de este curso
 Los estudiantes de administración de empresas de la UCR
 Los estudiantes de la UCR
 Los estudiantes de educación superior en Costa Rica
 Los trabajadores asalariados en Costa Rica

Población: conjunto total de las unidades de estudio (finita, infinita)

Finita: Es posible contar en número de miembros o el número de observaciones posibles o valores que puede tomar la característica.

Ejem:

- Cantidad de estudiantes en la UCR en el año 2009
- Número de hermanos de los estudiantes de este grupo
- Cantidad de goles del clásico Saprissa – Alajuela
- Cantidad de huevos que pone una gallina en un mes

Infinita: Los posibles valores que pueden tomar las observaciones es infinita o muy difícil de contar. (variables continuas)

Ejm.

- Peso de la producción de café en la cosecha 2008
- Tiempo que se tarda viajando en automóvil entre San Pedro y San Ramón

- Producción de leche en litros de una vaca en un mes.

Por asuntos matemáticos es importante esta diferenciación, particularmente es útil en matemáticas trabajar con variables continuas.

1.3 Fuentes de información

La información requerida para realizar una investigación puede estar disponible en algún lugar (datos existentes), Ejm. Los salarios de los funcionarios de la Sede Occidente. O es posible que haya que generar los datos (datos no disponibles), ejm. Percepción de los estudiantes sobre la calidad de los servicios que presta la Soda.

En el primer caso hablamos de *fuentes de información*, en el segundo caso hablamos de *métodos* para recolectar la información.

- *Fuente primaria*: es aquella persona, agencia y organización que proporciona o publica datos que ella misma ha generado. Ejm INEC, BCCR, MH.
- *Fuente secundaria*: se refiere a persona, agencia u organización que hace publicaciones, donde presenta información recopilada por otra fuente. Ejm. Los datos (sobre empleo) de una tesis que utilizó las publicaciones de INEC
 - Generalmente se prefieren las fuentes primarias porque:
 - Evita errores de transcripción,
 - Generalmente incluye más detalle.
 - Incluye la definición de los términos utilizados
 - Describe el procedimiento que se utilizó en la investigación

1.4 Métodos de recolección de información no existente (Quintana)

- *Observación*: El investigador recoge los datos observando lo que le interesa y anotando las observaciones. Se usa más en las ciencias físicas y naturales. Ejm. Observar el crecimiento diario de X variedad de maíz. Ventaja: ofrece información muy completa y confiable; desventaja: requiere de un diseño bien elaborado, y su aplicación generalmente implica mucho costo y tiempo.
- *Entrevista*. Generalmente a partir de un cuestionario, se visita al informante para obtener la información. Frecuente en ciencias sociales. Ejm. Un estudio de mercado para saber el consumo de X artículo. Ventaja, buena información; desventaja, alto costo.
 - Por Teléfono.

- *Correo*: Similar al anterior pero el cuestionario se envía por correo. Ventaja, más económico. Problema: baja respuesta y malas interpretaciones de las preguntas.
 - Por Internet, ágil, pero contesta el que quiere
- *Registro*: Se refiere más a información donde oficialmente se determina que determinada institución registre y suministre la información: ejm nacimientos, defunciones, matrimonios, comercio exterior, registro de capitales. Ventaja, económico; desventaja, posibles omisiones en los registros. Ejm. Las importaciones con otros nombres o de procedencias escritas con otros nombres. Caso concreto, las importaciones de Estados Unidos que a veces se escribe Estados Unidos, EU, EEUU, USA, etc.

1.5 Fases de una investigación estadística

- Definición del problema y planteamiento de objetivos
- Preparación del plan de trabajo
- Preparación y prueba del cuestionario
- Diseño y selección de la muestra
- Preparación y ejecución del trabajo de campo
- Procesamiento de la información
- Análisis e interpretación de los datos
- Preparación del informe final

1.6 Necesidad de trabajar con muestras. (si la población es muy grande)

- Más fácil
- Más barato
- A veces se pierde la población (semillas)
- Generalmente los resultados son similares (sangre)
- Tamaño: 5% de población, dependiendo de homogeneidad

1.7 Diseño y selección de la muestra, tipos de muestreo.

- *Por conveniencia*: Se seleccionan los más convenientes (fáciles)
- *Intencional*, utilizando el juicio de un experto.
- *Muestreo aleatorio* (al azar). Todas las unidades tienen igual posibilidad
 - Muestreo simple al azar: Enumero y luego se selecciona la muestra utilizando rifa, tabla de número o excel.
 - Muestreo sistemático: como el del AID; si $N = 1000$ y $n = 100$, $N/n = 10$. Se selecciona al azar un número entre 1 y 10, los restantes 99 se seleccionan sumando de 10 en 10 a partir del primer número elegido.
 - Muestreo estratificado: por grupos. Ejm en un estudio de intención de voto en la UCR, separar profesores, administrativo y estudiantes; luego elegir al azar una muestra representativa de cada estrato.

- Muestreo por conglomerados: Ejm por provincias
- Ejm. Si quiero estudiar el rendimiento académico en la Sede Occidente, por conveniencia selecciono solo a este grupo, intencional, le pregunto al director cuál grupo es más homogéneo, al azar con el sistema de lotería.
- Cómo hacer la selección al azar? Números al azar, tablas y Excel

1.8 Generación de números al azar con EXCEL

Trabajar con el comando de funciones f(x), aleatorio (random)

Definir generar un número al azar en determinado rango

Fórmula: aleatorio()* [(>-<)+<]. De 0 a 100 sería aleatorio()*100. F9 fija a valores

Luego copiar esa celda la cantidad de veces requerida (para generar esa cantidad de números aleatorios).

Copiar en otra hoja con “pegado especial” valores.

2. números relativos

El análisis estadístico se basa, generalmente, en información numérica, cuantitativa (sin dejar de lado la parte cualitativa).

En ese sentido es importante (pero insuficiente) el estudio de *valores absolutos*.

Ejm.1 La cantidad de estudiantes que piden admisión a la Sede San Ramón, para el periodo 2009. Ello permite determinar la cantidad de nuevos grupos, número de estudiantes por grupo, cantidad de profesores requeridos, etc.

Ejm. 2 Cantidad de vehículos Toyota Yaris vendidos en 2008. Esa información permite programar las importaciones de esos autos para el 2009.

Ejm. 3. Cantidad de créditos solicitados al Banco Nacional en el 2008.

EJm. 4 Déficit fiscal de Costa Rica en el año 2008.

Sin embargo a veces el manejo de magnitudes absolutas es insuficiente, porque:

- Las magnitudes a veces son muy grandes y difícil de recordar (E, exp. Df)
- Interesa hacer relaciones entre variables, o su evolución en el tiempo.
- Conviene hacer comparaciones entre grupos de datos.

Para facilitar el análisis cuantitativo, se recurre al uso de “*números relativos*”, que son expresiones que permiten simplificar el comportamiento de una variable determinada y enriquecer el análisis.

Entre estas expresiones se pueden citar: Razones, proporciones, porcentajes, índices.

2.1 Razones

Es la relación entre dos números positivos, que pueden o no pertenecer al mismo universo. Si A y B son números positivos, una razón sería A/B que indica cuantas veces cabe A entre B.

Ejm, si A = total estudiantes de la UCR en 2008 = 30 mil

B = total estudiantes de ITC en 2008 = 10 mil

A/B = 30,000/10,000 = 3

Indica que hay 3 estudiantes en la UCR por cada estudiante en el ITCR.

Ejm, comparar la población de Costa Rica con la de Estados Unidos.

Ejm. Cantidad de mujeres respecto a los hombres en este grupo

Las razones pueden tomar cualquier valor positivo, como 4 o 0.2.

2.2 Proporciones

Son un tipo particular de razones, que cumplen dos características:

- Relaciones dos números del mismo universo
- Relacionan una parte con el todo

Algebraicamente, sean tres números que pertenecen al mismo universo A B, C; una proporción sería $A / (A+B+C)$, que indica cuánto representa A del total del universo, compuesto por la sumatoria de los tres subgrupos A, B, y C.

Otras proporciones serían: B/T , C/T , $(A+B)/T$, $(A+C)/T$, $(B+C)/T$, T/T

Dada esas características, el valor de una proporción siempre varía entre 0 y 1.

Ejm. Proporción de hombres en este grupo:

A = número de hombres (2)

B = número de mujeres (3)

$A/A+B = 2/5 = 0.4$.

Ejm. En tres grupos de estudiantes $A= 20$, $B = 30$, $C = 50$. Calcular las proporciones.

2.3 Porcentajes

Generalmente los resultados de las razones constituyen números quebrados, que son incómodos de manejar, por lo que se recurre a llevarlos a una base más amigable. La base más utilizada es 100, que implica multiplicar el resultado de la razón por 100.

Ejm. Si en el ejercicio anterior el resultado se multiplica por 100, se obtiene:

$0.4*100 = 40\%$.

Hay que tener cuidado con la interpretación de los números cuando hay *cambio de base*.

Ejm. Si partimos de una cantidad = 500 que se reduce en 20%, tenemos 400.

Para volver de 400 a 500 no se logra incrementando 400 en 20% (80). Hay que aumentarlo en $100/400 = 25\%$.

La variación absoluta es la misma, pero como corresponde a diferente cantidad base, el % de cambio es diferente en cada caso, se requiere un cambio porcentual más grande cuando la base es menor para obtener la misma magnitud.

Ejercicio:

Si 10,000 se reduce en un 50%. Qué valor se obtiene y en qué porcentaje se debe incrementar ese valor para llegar nuevamente a 10,000?

2.4 Tasas

Las tasas son un indicador que a veces se utiliza incorrectamente, como porcentaje.

Este concepto hace referencia más bien a variables que cambian en el tiempo.

Generalmente se refiere a la variable población, cuya *base* crece cada año.

Ejm, si la población de determinado país es de 4 millones y en un periodo de 10 años pasó a 5 millones, tuvo una tasa de crecimiento de un 25% (un millón de habitantes); si en los siguientes 10 años creció también un 25%, pasó a 6.25 millones (creció 1.25 millones).

Como la base es diferente (4, 5 millones) un 25% en cada caso equivale a montos absolutos diferentes (1 millón y 1.25 millones, respectivamente).

$$\text{Tasa: } [(t_2 - t_1) / t_1 - 1] * 100 = [(t_2 / t_1) - 1] * 100$$

El caso clásico de tasas, son las tasas de interés. Interés simple y compuesto.

2.4.1 Interés simple

Si tengo un millón de colones y lo coloco al 10% anual por cinco años, al final de cada año puedo retirar 100 mil colones. Si M_0 =principal, M_r =rendimiento;

$$M_1 = M_0 + M_r = M_0(1+r)$$

$$M_0 = 1,000,000 = \text{monto inicial}$$

$$M_1 = M_0(1+r) = 1,000,000*(1.1) = 1,100,000$$

$$M_2 = M_0(1+r+r) = M_0(1+2r) = 1,000,000*(1.2) = 1,200,000$$

$$M_3 = M_0(1+r+r+r) = M_0(1+3r) = 1,000,000*(1.3) = 1,300,000$$

$$M_4 = M_0(1+r+r+r+r) = M_0(1+4r) = 1,000,000*(1.4) = 1,400,000$$

$$M_5 = M_0(1+r+r+r+r+r) = M_0(1+5r) = 1,000,000*(1.5) = 1,500,000$$

$$\rightarrow M_t = M_0 (1+tr)$$

Ejercicio:

Si tengo 200,000 y los coloco al 5% por 12 años, ¿Cuánto tengo al final de los 12 años (principal más intereses generados, asumiendo que cada año retiro los intereses)?

$$R/200,000(1+12*0.05) = 320,000$$

El tiempo cuenta en la obtención del dinero, lo prefiero hoy que mañana.

Estaría dispuesto a dejar los intereses hasta el final solo si gana intereses sobre intereses.

2.4.2 Interés compuesto

Este concepto se refiere a aquel caso donde la base cambia en cada periodo. Este comportamiento se puede representar algebraicamente de la siguiente manera:

$$M1 = Mo(1+r) = 1,000,000*(1.1) = 1,100,000; \text{ si dejo en el banco } M1 \Rightarrow$$

$$M2 = M1*(1+r)$$

$$M2 = Mo(1+r)*(1+r) = Mo*(1+r)^2$$

$$M3 = M2(1+r) = Mo(1+r)^2(1+r) = Mo(1+r)^3$$

$$M4 = Mo(1+r)^4$$

$$M5 = Mo(1+r)^5$$

$$Mo = 1,000,000$$

$$M1 = M1(1+r) = 1,100,000$$

$$M2 = M2(1+.1)^2 = 1,210,000$$

$$M3 = M2(1+.1)^3 = 1,331,000$$

$$M4 = M2(1+.1)^4 = 1,464,100$$

$$M5 = M2(1+.1)^5 = 1,610,510$$

$$\rightarrow Mt = Mo(1+r)^t$$

Ejercicio.

Si tengo 200,000 y los coloco al 5% por 12 años, y retiro los intereses y el principal al final, ¿cuánto tendré al cabo de 12 años?

$$R/ 200,000(1.05)^{12} = 359,171. \quad \text{Comparar con la anterior fórmula.}$$

A veces lo que se tiene es el monto actual y el que recibirá al final de un periodo determinado, pero no se conoce la tasa, la cual debe ser calculada. Ello se logra despejando r en la fórmula anterior:

$$Mt = Mo(1+r)^t \Rightarrow$$

$$Mt/Mo = (1+r)^t$$

$$(Mt/Mo)^{(1/t)} = 1+r$$

$$(Mt/Mo)^{(1/t)} - 1 = r$$

Con los datos del ejemplo anterior se tiene:

$$(1,610,510/1,000,000)^{(1/5)} - 1 = r = .1 = 10\%$$

Alguna gente vendió una finca a pagos y recibe pagarés, que puede descontar. En este caso la variable que no se conoce es Mo , la cual se puede calcular considerando el valor actual de r y despejando Mo en la fórmula anterior, así:

$$M_t = M_o(1+r)^t$$

$$M_t/(1+r)^t = M_o$$

En el ejemplo anterior:

$$1,610,510/(1.1)^5 = 1,000,000$$

Ejercicio.

Su jefe le dijo que si seguía con el mismo desempeño, le daría una bonificación de un millón de colones dentro de 3 años. ¿Cuál es el valor actual de esa bonificación?

R/

$$M_o = M_t/(1+r)^t = 1,000,000/(1.05)^t = 863,837.7$$

PRACTICA

Juan hizo un depósito por 2 millones de colones en el banco nacional, a una tasa del 12% anual por 5 años, y recibirá el total al final del periodo. ¿cuánto obtendrá Juan al final de los 5 años?

R/

$$M_t = M_o(1+r)^t$$

$$M_5 = 2,000,000(1.12)^5 = 2,000,000(1.76234) = 3524,683$$

Diana tiene un pagaré por 10 millones que vence dentro de 5 años. Si la tasa de interés promedio en el mercado es de 10%. ¿cuánto vale ese pagaré hoy?

R/

$$M_5 = M_o(1+r)^t \Rightarrow \Rightarrow$$

$$M_5 / (1+r)^t = M_o$$

$$10,000,000/(1.1)^5 = 6,209,213$$

Fernando compró hoy un título valor en un millón de colones, que le paga 2 millones dentro de 10 años. ¿cuál es la tasa de rendimiento de esa inversión?

R/

$$M_t = M_o(1+r)^t$$

$$(M_t/M_o)^{(1/t)} - 1 = r$$

$$(2/1)^{(1/10)} - 1 = 7.17\%$$

2.5 Los números índices

Los índices son indicadores estadísticos muy utilizados, particularmente cuando interesa analizar la tendencia de una variable determinada, más que su magnitud absoluta, que cuando los números son grandes se dificulta su manejo.

Ejem, un índice que muestra la tendencia de la flota vehicular en el país.

año	No. de vehículos	Índice $= (\text{Año}_i / \text{año}_o) * 100$
2000	400,000	100
2001	480,000	120
2002	500,000	125
2003	600,000	150
2004	800,000	200

Metodología:

Se define un año base.

Se calcula el valor del índice así: $= (\text{Año}_i / \text{año}_o) * 100$

Para el año base el valor del índice es 100

Luego se pueden calcular tasas de crecimiento de la variable a partir del índice.

Para el año 2002, que el índice es 125, indica que del año 2000 al 2002 la flota vehicular del país creció en un 25% (125-100).

El índice puede tener cualquier año base. Si se toma como año base el último de la serie, entonces el valor del índice en vez de crecer, decrece a partir de 100, a medida que nos alejamos en el tiempo hacia atrás.

2.5.1 Índices de precios

Quizás los índices más comúnmente utilizados son los índices de precios. Existen distintos de éstos:

Índice de precios al consumidor

Índice de precios al productor

Índice de precios de los servicios

Índice de precios de la construcción, etc.

Fórmula de Laspeyres: $I = \frac{\sum P_n Q_o}{\sum P_o Q_o}$

Fórmula de Paasche: $I = \frac{\sum P_n Q_n}{\sum P_o Q_n}$

2.5.1.1 El IPC en Costa Rica

Labor del INEC. Para calcular el IPC el INEC debe:

- Definir el año (mes) base

- Elegir los bienes a ser incluidos: 292, con base en encuesta 2004
- Definir las ponderaciones para cada artículo
- Elección de la fórmula del índice (combinación de ambas)
- Definir el procedimiento de recolección de la información

Limitación del IPC: sólo incluye algunos bienes.

2.5.1.2 Usos del IPC

El IPC refleja la variación en el poder adquisitivo de la moneda del país. A medida que crece el IPC, se reduce el poder de compra.

Esta característica hace al IP un indicador útil para ajustar algunas variables como salarios, presupuestos, ingresos fiscales y otros. Para calcular la evolución del poder adquisitivo se usan los valores “reales” o “deflatados”.

Valor real en año_t = valor nominal en año_t/IPC_t* 100

Ejm. Diana ganaba 400,000 colones en diciembre 2008. ¿cuál era su salario real, tomando como base julio 2006?

R/

$$SR = sn/IPC*100 = 400,000/129.95 = 307,810$$

Pero más importante que el salario real (SR) aislado en un periodo dado, conviene analizar su evolución.

Ejm José Pablo ganaba 500 mil colones en julio 2006 y 600 mil en diciembre 2008. Cuándo tenía José Pablo un mayor poder adquisitivo, según el IPC que calcula el INEC para Costa Rica?

R/

$$SR \text{ en año}_t = SN \text{ año}_t / IPC_t * 100$$

$$SR_{dic.2008} = (600,000/129.95)*100 = 461,716$$

José Pablo tenía un mayor poder adquisitivo en julio 2006

Para un periodo determinado la tasa de crecimiento real se calcula así:

$$\text{Tasa de crecimiento real} = [(1+\Delta PIB)/(1+\Delta IPC) - 1] * 100$$

Ejm. Si en determinado periodo el salario monetario de Fernando creció un 25% y el IPC creció en 20%, el SR de Fernando creció en un 4.16%. Esto se calcula así:

$$\Delta SR = (1.25/1.20) - 1 * 100 = 4.16\%.$$

Si solo tiene los datos absolutos del salario de Fernando, debe calcular su tasa de crecimiento así:

$$(T_2 - t_1) / t_1 * 100.$$

Ejercicio.

Las ventas anuales (enero-diciembre) de la empresa XY fueron las siguientes:

Año	Ventas en \$	IPC _{base j.2006}
2004	1,000,000	82.48
2005	1,100,000	94.09
2006	1,300,000	102.96
2007	1,400,000	114.09
2008	1,600,000	129.95

- a) Calcule las ventas en términos reales de la empresa XY para los años respectivos,
- b) Calcule la tasa de crecimiento real de las ventas en los años:
 - i. 2005-2006,
 - ii. 2006-2007
 - iii. 2007-2008
 - iv. 2004-2008
- c) Comente los resultados obtenidos

R/

a)

año	Ventas en \$	IPC _{base j.2006}	Ventas reales
2004	1,000,000	82.48	1,212,415
2005	1,100,000	94.09	1,169,093
2006	1,300,000	102.96	1,262,626
2007	1,400,000	114.09	1,227,101
2008	1,600,000	129.95	1,231,242

b) Calcule la tasa de crecimiento real de las ventas en los años:

- i. 2005-2006 = -3.57%
 - ii. 2006-2007 = -2.81%
 - iii. 2007-2008 = 0.33%
 - iv. 2004-2008 = 1.55%
- d) Aunque las ventas nominales crecen mucho, las reales casi no crecen.

Universidad de Costa Rica
Sede Occidente
Departamento de Ciencias Sociales
Curso: Estadística general 1
Prof. Carlos Conejo
Periodo: verano 2009.

Práctica sobre uso del IPC

Las ventas anuales (enero-diciembre) de la empresa XY fueron las siguientes:

año	Ventas en \$	IPC _{base j.2006}
2004	1,000,000	82.48
2005	1,100,000	94.09
2006	1,300,000	102.96
2007	1,400,000	114.09
2008	1,600,000	129.95

Con base en la información anterior se pide:

- e) Calcule las ventas en términos reales de la empresa XY para los años respectivos
- f) Calcule la tasa de crecimiento real de las ventas en los años:
 - i. 2005-2006,
 - ii. 2006-2007
 - iii. 2007-2008
 - iv. 2004-2008
- g) Comente los resultados obtenidos

Universidad de Costa Rica
Sede Occidente
Departamento de Ciencias Sociales
Curso: Estadística general 1
Prof. Carlos Conejo
Periodo: verano 2009.

Práctica 2

1. En diciembre de 2008, ante la crisis económica, la empresa Toyota de Costa Rica decidió reducir en un 30% las comisiones por ventas que paga a sus vendedores. Además decidió aumentarlas paulatinamente, un 10% el II trimestre de 2009, un 10% en III trimestre y otro 10% en el IV trimestre. Ante esta medida el Gerente de la Toyota indicó que “así en el IV trimestre del año 2009 las comisiones alcanzarán el nivel que tenían antes de la modificación. Comente la afirmación del Gerente.
2. Según el INEC en el cantón de San Ramón la Tasa Bruta de Natalidad fue de 20 por mil en el año 2006. Si a mediados de ese año la población de ese cantón rondaba los 220,000 habitantes. ¿aproximadamente cuántos nacimientos ocurrieron en San Ramón en el año 2006?
3. El director de una escuela privada desea calcular el ingreso que tendrá en el año 2009 y 2010 por concepto de matrícula y mensualidad que pagan los estudiantes. El año pasado ese monto fue de 20 millones de colones y la tasa de crecimiento promedio en los últimos años ha sido de 10%. Calcule el monto esperado de ingresos para los años 2009 y 2010, asumiendo que la tasa de crecimiento se mantiene en 10%.
4. Asuma que Costa Rica tiene una población de 4 millones de habitantes y que crece a una tasa del 2.5% anual. ¿Cuál será la población de Costa Rica dentro de 10 años?
5. En enero del año 2007, Karol ganaba 800 mil colones, y en agosto del 2008 ganaba un millón de colones. Para esos meses el IPC fue de 103.95 y 126.38, respectivamente. Calcule la variación en el poder adquisitivo del salario de Karol.
6. Un dirigente sindical indicó a un noticiero que “es necesario incrementar el número de artículos que componen la canasta del IPC, así la medición de la inflación sería mayor reflejando la realidad del país y ese nuevo instrumento nos permitirá justificar mayores incrementos salariales”. Comente esa afirmación.

3. Distribuciones de frecuencia

Vimos que los datos estadísticos generalmente provienen de observaciones de una característica de determinada población (unidad estadística). Cuando el número de observaciones es pequeño, se puede (y debe) hacer el análisis refiriéndose a todos y cada uno de los elementos.

Por ejemplo, si al final de este curso quiero hacer un análisis del rendimiento académico de este grupo (10 estudiantes) puedo hacerlo indicando el comportamiento de cada uno. Específicamente, podría decir, los 3 mejores estudiantes ganaron el curso con una nota de 100, 2 retiraron el curso y su nota es de 0. De los 5 restantes, 2 pasaron con un 80 y tres con un 90. Después puedo sacar la nota promedio, resaltar quien tiene la nota más alta y quien la más baja, etc.

Pero en la práctica generalmente se trabaja con un número mayor de datos, a veces cientos o incluso miles, en cuyo caso sería muy tedioso referirse a cada uno de los resultados. . Aún con grupos más pequeños, como 40 o 60 observaciones resulta difícil y tedioso el análisis sin resumir la información.

Aquí es donde surge la necesidad o conveniencia de *resumir* (agrupar) la información, o de establecer una *distribución de frecuencia (grupos)*, para simplificar y facilitar el análisis

En el análisis e interpretación de una serie de datos es importante tener información sobre tres aspectos, a saber:

- La *forma* de distribución de los datos (estructura de la distribución)
- La *posición* de la distribución (valores centrales)
- La *dispersión* de los datos respecto al valor central (variabilidad)

Seguidamente nos referiremos al primer aspecto y en las clases siguientes estudiaremos los otros dos temas.

definición: Una distribución de frecuencia es una ordenación o arreglo de datos en clases o categorías (grupos) que muestran, para cada una de ellas, el número de elementos que contienen o frecuencia.

Variables ordenadas.

Aquí pueden existir variables discretas o continuas.

En las continuas está el asunto de los límites: reales e indicados (redondeo).

Pero Excel lo soluciona (Desde $>$ que el límite inferior hasta \leq al límite superior).

Definición de las clases: entre 6 y 15, dependiendo de los datos.

Ejem. De las notas de un grupo de 25 estudiantes:

Clase	Frecuencia
0-20	2
21-40	3
41-60	5
61-80	4
81-100	<u>11</u>
Total	25

Entregar y discutir copia de datos originales, ordenados y la distribución.

Definición de límites. EXCEL

EJEMPLO. Pág. 245.

Frecuencia absoluta y relativa, simple y acumulada

(Fotocopiar el cuadro pág. 252 para explicar esto)

Hablar de graficar.

- Histogramas (gráficos de barras)
- Polígonos de frecuencia (gráficos de líneas)
- Ojivas (gráfico de frecuencia acumulada)

Prácticas 3 y 4.

Luego indicar que eso se simplifica con Excel.

Distribución de Frecuencia

(peso en kilogramos de 40 estudiantes)

49	60	46	37	54
43	59	40	62	47
46	52	55	41	66
45	36	50	51	48
42	53	53	68	56
35	74	43	47	60
51	67	40	70	49
41	46	32	57	43

(datos ordenados de manera ascendente)

32	42	47	52	60
35	43	47	53	60
36	43	48	53	62
37	43	49	54	66
40	45	49	55	67
40	46	50	56	68
41	46	51	57	70
41	46	51	59	74

(Datos distribuidos en frecuencias)

Clase (peso en Kgrs.)	Frecuencia (fi)
30-34	1
35-39	3
40-44	8
45-49	9
50-54	7
55-59	4
60-64	3
65-69	3
70-74	<u>2</u>
	Total
	40

Cuadro 7.2:
**CLASIFICACION DE LOS 40 ESTUDIANTES DE UN COLEGIO
 DE ACUERDO A SU PESO EN KILOGRAMOS**

CLASES	PUNTOS MEDIOS x_i	FRECUENCIA		ACUMULADA "menos de"		ACUMULADA "más de"	
		Absoluta f_i	Relativa f_r	Absoluta	Relativa	Absoluta	Relativa
29,5 - 34,5	32	1	0,025	1	0,025	40	1,000
34,5 - 39,5	37	3	0,075	4	0,100	39	0,975
39,5 - 44,5	42	8	0,200	12	0,300	36	0,900
44,5 - 49,5	47	9	0,225	21	0,525	28	0,700
49,5 - 54,5	52	7	0,175	28	0,700	19	0,475
54,5 - 59,5	57	4	0,100	32	0,800	12	0,300
59,5 - 64,5	62	3	0,075	35	0,875	8	0,200
64,5 - 69,5	67	3	0,075	38	0,950	5	0,125
69,5 - 74,5	72	2	0,050	40	1,000	2	0,050
TOTAL		40	1,000				

4. Medidas de posición

Vimos que cuando analizamos el comportamiento de una variable generalmente contamos con muchas observaciones, que es necesario simplificar de alguna forma para facilitar la interpretación de los datos. Las distribuciones de frecuencia constituyen una excelente ayuda en este aspecto.

Sin embargo, a veces conviene tener “una medida de tendencia central”, un solo indicador que nos de una idea que resuma el nivel de la variable que estamos analizando. Para ello se han definido principalmente 3 medidas de tendencia central. Que son las siguientes:

- Moda
- Mediana
- Media aritmética (promedio)

La moda (Mo)

Es un término muy conocido, que se utiliza frecuentemente. Tiene la ventaja que también se puede utilizar cuando se analiza una variable cualitativa.

Ejm. Cuáles son los celulares que están de moda.Cuál es el color de moda

Def. es el valor que más veces aparece en una serie de observaciones

Una serie puede ser unimodal, bimodal, multimodal, sin moda.

Ejm. En la serie: 12233344455555666777889, moda= 5

La mediana (Me)

Definición: La mediana constituye el valor central de una serie de datos ordenados. En otras palabras, el 50% de los valores de la serie son \leq Me y el 50% son \geq Me.

Posición de la mediana = $(n+1)/2$

Ejm: sea la serie: 6, 8, 8, 10, 12, 19, 23

Posición de la Me = $(n+1)/2 = (7+1)/2 = 4$

Entonces vemos en la serie que Me = 10

En series con $n = \text{par}$, la posición de la Me queda entre dos números.

La media aritmética o promedio (pr)

Es la medida de tendencia central más conocida y más utilizada, además la más útil estadísticamente hablando.

Pr = suma de los valores de las observaciones/ número de observaciones = $\sum Xi/n$

Ejm: el ingreso en miles de 3 profesores = 100, 200 y 300:

Pr = $(100+200+300)/3 = 200$ mil

El pr es aquel número que hace que la suma de las desviaciones de las observaciones respecto al pr sea cero.

El promedio se ve afectado por los valores extremos (diferente a la Mo y la Me).

Ejm. En la serie: 1,2,2,3,3,3,4,4,4,5,5,5,5,5,6,6,6,7,7,7,8,8,9, moda=5, Me=5, Pr=5 pero si la modificamos: 1,2,2,3,3,3,4,4,4,5,5,5,5,5,6,6,6,7,7,7,8,8,20, Mo=5; Me=5, pr=5.44.

Promedio ponderado

Algunas veces se cuenta con datos que tienen distinto peso relativo, de manera que se debe considerar esa condición para calcular adecuadamente el promedio. Eso se llama promedio ponderado porque “pondera” cada componente respecto a su importancia.

El ejemplo más cercano son las notas de un determinado curso:

Se tienen los siguientes ítems para determinar la calificación en un curso de administración:

Item	Porcentaje
Primer examen parcial	25%
Segundo examen parcial	25%
Participación	10%
Trabajo final	40%

Si todos los ítems valen igual (25%), y las calificaciones de Karol son:

I parcial 65

II parcial 75
 Participación 80
 Trabajo final 60

El promedio sería: $\sum Xi/n = (65+75+80+60)/4 = 280/4 = 70$, ganó el curso.

Pero si los ítems tienen diferente ponderación, entonces:

Item	porcentaje	Nota	Nota*porcentaje
I parcial	25%	65	1625
II parcial	25%	75	1875
participación	10%	80	800
Trabajo final	40%	60	2400
Total	100	280	6700

$$\text{Pr.p.} = (\sum xi * P) / \sum P = 6700 / 100 = 67$$

Ejemplo.

Se tienen tres categorías de trabajadores con sus respectivos salarios, así:

Categoría	Cantidad de trabajadores	Salario en miles de ¢
Docentes	10	400
investigadores	20	500
Administrativos	40	300

$$\text{Pr} = : \sum Xi/n = 1200/3 = 400$$

$$\text{Pr.P} = (\sum xi * P) / \sum P = \phi$$

Categoría	No. trabajadores	Salario	N*P
Docentes	10	400	4,000
Investigadores	20	500	10,000
Administrativos	40	300	12,000
Total	70	1,200	15,200

$$\text{Prom.} = \sum xi/n = 1200/3 = 400$$

$$\text{Pr.P} = (\sum xi * P) / \sum P = 26,000/70 = 371$$

Ejercicio

Costa Rica es un país que cuenta con tres regiones, cuya población y consumo per cápita de carne en kilogramos es el siguiente:

Región	población (P)	Consumo de carne	P*Kgrs
A	200,000	4	800,000
B	300,000	6	1,800,000
C	500,000	8	4.000,000
Total	$\sum P = 1,000,000$	$\sum x_i = 18$	$\sum x_i * P = 6,600,000$

Con base en la información anterior, calcule el promedio ponderado de consumo de carne.

R/

$$\text{Promedio simple} = \sum x_i / n = (4+6+8)/3 = 6$$

$$\text{Promedio ponderado} = (\sum x_i * P) / \sum P = 6,600,000 / 1,000,000 = 6.6.$$

Si la población total de Costa Linda es de dos millones ¿Cuál es el consumo de carne?

$$\text{R/ } 2,000,000 * 6.6 = 13,200,000$$