

UTILIDAD DE LA ESTADÍSTICA PARA LOS NEGOCIOS Y LA ECONOMÍA (15 MINUTOS)

La **estadística** consiste en un grupo de instrumentos, herramientas y técnicas que nos ayudan a describir, organizar, recolectar e interpretar datos. Pero eso no significa que estemos dando vueltas en números y formulas indescifrables. Como dice la frase cliché, la estadística está en todas partes. Aunque no nos demos cuenta, en cada cosa que hacemos, ya sea por diversión, estudio o trabajo estamos utilizando herramientas estadísticas. Al jugar una polla en la Copa América o apostar al goleador del campeonato. Al calcular qué nota necesitamos en el examen para aprobar el ramo y al decidir qué ejercicio dejo para el final. Al decidir dónde me conviene invertir según la rentabilidad de las acciones o al pedir un préstamo. También cuando pensamos en qué lugar nos conviene almorzar según el precio, calidad y ubicación del local. E incluso si salgo abrigado o no dependiendo de cómo estuvo el clima en los días anteriores... o quizás confío en Iván Torres con cierto riesgo e incertidumbre.

Este **curso** los acercará a la estadística con ejemplos del mundo que los rodea, con noticias del momento y situaciones cotidianas de su vida actual y su futura vida como ingenieros comerciales o economistas. Las estadísticas son necesarias para la vida cotidiana y para muchas otras áreas del conocimiento, ya sean de tipo técnico o científico.

Entre las **funciones** de la estadística se encuentran recolectar, organizar, presentar, analizar e interpretar datos:

- Recolectar datos por medio de encuestas y mediciones de distintos tipos.
- Organizar datos recolectados o datos ya trabajados.
- Presentar los datos ordenados de manera lógica y describirlos de tal forma que revelen de manera sencilla la información que contienen.
- Analizar e interpretar datos implica evaluar su calidad, posibles sesgos, incertidumbre, controlar por distintos factores, comprender las relaciones observadas y no observadas, identificar efectos, evaluar Impacto.

Las **aplicaciones** de las estadísticas también son variadas e incluyen, entre otras:

- **Contabilidad:** Tomar decisiones sobre clientes de confianza, los que pagan a tiempo, los con deudas y los que generan pérdidas. Para diferenciar ventas dentro de la empresa, determinar cuáles productos son aceptados y cuáles no. Ejemplo: determinar si las cuentas por cobrar en la hoja de balance de un cliente representan la verdadera cantidad de cuentas por cobrar.
- **Finanzas:** Tomar decisiones sobre inversiones. Ejemplos: crear carteras de inversión y calcular la relación precio/ganancia o el rendimiento de dividendos.
- **Marketing:** Tomar decisiones sobre lanzamientos de nuevos productos, reposicionamiento y renovación. Ejemplo: para desarrollar futuras estrategias de marketing es útil contar con resúmenes estadísticos sobre actividades promocionales, precios y posicionamiento.
- **Industria:** Vigilancia de la producción, controles de calidad, tiempos de transporte y almacenamiento. Ejemplo: en la producción de un nuevo filamento para focos eléctricos se puede hacer un control de calidad utilizando una muestra para determinar la duración promedio de los nuevos focos.
- **Economía:** Análisis de datos, investigación, distintos pronósticos sobre diversos aspectos de la economía. Ejemplo: calcular la tasa de inflación, proyectar el desempleo, evaluar compra de bonos, determinar la capacidad de producción.

ESTADÍSTICA DESCRIPTIVA VERSUS ESTADÍSTICA INFERENCIAL (45 MINUTOS)

En la malla de ICO, este curso de estadística inferencial se ubica a continuación de estadística descriptiva:

- **Estadística descriptiva:** Conjunto de procedimientos y herramientas que tienen como objetivo sintetizar datos en tablas, gráficos y medidas de resumen. Suele ser la primera fase en el análisis de información.

- *Estadística inferencial*: Conjunto de procedimientos y herramientas que tienen como objetivo hacer estimaciones y probar hipótesis sobre las características de una población, a partir de una muestra.

Ejemplo de Estadística Descriptiva

Las notas de las pruebas de matemáticas de 1.000 alumnos de enseñanza media de un determinado colegio pueden ser descritas/resumidas de la siguiente forma:

- *Frecuencia absoluta*: Número de repeticiones de una observación.
- *Frecuencia relativa*: Frecuencia absoluta dividida por el número total de datos.
- *Frecuencia absoluta acumulada*: Suma de los distintos valores de la frecuencia absoluta tomando como referencia un individuo dado. Último valor es igual al número de casos.
- *Frecuencia relativa acumulada*: Frecuencia absoluta acumulada dividida por el número total de datos.

NOTA	FRECUENCIA			
	ABSOLUTA	RELATIVA	ABSOLUTA ACUMULADA	RELATIVA ACUMULADA
1,2	1	0%	1	0%
1,4	2	0%	3	0%
1,6	3	0%	6	1%
1,8	8	1%	14	1%
2,0	14	1%	28	3%
2,2	18	2%	46	5%
2,4	19	2%	65	7%
2,6	22	2%	87	9%
2,8	25	3%	112	11%
3,0	26	3%	138	14%
3,2	27	3%	165	17%
3,4	31	3%	196	20%
3,6	35	4%	231	23%
3,8	38	4%	269	27%
4,0	45	5%	314	31%
4,2	46	5%	360	36%
4,4	48	5%	408	41%
4,6	52	5%	460	46%
4,8	58	6%	518	52%
5,0	60	6%	578	58%
5,2	56	6%	634	63%
5,4	54	5%	688	69%
5,6	51	5%	739	74%
5,8	50	5%	789	79%
6,0	46	5%	835	84%
6,2	44	4%	879	88%
6,4	40	4%	919	92%
6,6	32	3%	951	95%
6,8	31	3%	982	98%
7,0	18	2%	1000	100%
TOTAL	1000	100%		

Ejemplo de Estadística Inferencial: Chilectra S.A.

Chilectra S.A. está probando nuevos focos de luz entre sus más de 1.000 luminarias que monitorea y mantiene en la comuna de Santiago. La empresa quiere saber cuál es la vida útil de estos nuevos focos de luz en la comuna de Santiago.

Como la *población* de luminarias incluye a todos los focos de luz en la comuna, Chilectra decide hacer un estudio tomando una *muestra* de 200 focos de luz.

En la muestra de 200 focos, el promedio de vida útil de los focos es de 76 horas. Este resultado muestral sirve para estimar el promedio de la vida útil poblacional de los nuevos focos de luz instalados.

107	73	68	97	76	79	94	59	98	57
54	65	71	70	84	88	62	61	79	98
66	62	79	86	68	74	61	82	65	98
62	116	65	88	64	79	78	79	77	86
74	85	73	80	68	78	89	72	58	69
92	78	88	77	103	88	63	68	88	81
75	90	62	89	71	71	74	70	74	70
65	81	75	62	94	71	85	84	83	63
81	62	79	83	93	61	65	62	92	65
83	70	70	81	77	72	84	67	59	58
78	66	66	94	77	63	66	75	68	76
90	78	71	101	78	43	59	67	61	71
96	75	64	76	72	77	74	65	82	86
66	86	96	89	81	71	85	99	59	92
68	72	77	60	87	84	75	77	51	45
85	67	87	80	84	93	69	76	89	75
83	68	72	67	92	89	82	96	77	102
74	91	76	83	66	68	61	73	72	76
73	77	79	94	63	59	62	71	81	65
73	63	63	89	82	64	85	92	64	73

Discusión de Noticias Publicadas en la Prensa

- Noticia del diario La Tercera. 10 de Agosto, 2015. Encuesta realizada por CADEM (Empresa de investigación de mercado y opinión pública).
- Aprobación presidencial es medida con la siguiente pregunta: “Independiente de su posición política ¿Usted aprueba o desaprueba la forma como Michelle Bachelet está conduciendo su gobierno?”
- Tamaño de la muestra: 716 personas. 522 entrevistas telefónicas y 194 entrevistas presenciales.
- Margen de error de +/- 3,7 puntos porcentuales al 95% de confianza.
- ¿Creen ustedes que estos datos reflejan lo que opinaban los Chilenos? ¿Por qué?



Ejemplo de Estadística Descriptiva: Ranking de Universidades Chilenas

- Noticia diario Las Últimas noticias. 4 de Agosto, 2015. "Panorama de la Educación Superior en Chile 2014." Mineduc.
- 3 de cada 10 alumnos que ingresan a la educación superior abandonan su carrera antes de finalizar el primer año.
- ¿Creen ustedes que este promedio refleja el nivel de deserción que existía en el país? ¿Por qué?
- ¿En qué se diferencian las dos noticias discutidas?
- ¿Son ejemplos de estadística descriptiva o inferencial? ¿Por qué?

Las Últimas Noticias / Martes 4 de agosto de 2015

EMPLEO Y EDUCACIÓN

En algunas, más del 50% del alumnado abandona al primer año

Sorpresas en el ranking de deserción de las 59 universidades chilenas

ÓSCAR VALENZUELA

De cada 10 alumnos que ingresan a la educación superior, 3 abandonan su carrera antes de finalizar el primer año. Este promedio, extraído del "Panorama de la Educación Superior en Chile 2014" del Mineduc, pone de manifiesto el nivel de deserción que existe en el país.

Dejar botados los estudios es una tendencia mayor en Institutos Profesionales y Centros de Formación Técnica, pero también repercute con mucha fuerza en las universidades. De hecho, en la mayoría de

50% de deserción. En otras palabras, por insólito que parezca, la mitad de sus alumnos de primer año no llega a segundo.

Gabriel Ulloa, director de la escuela de Educación de la U. Bolivariana, identifica 3 razones que explicarían el temprano abandono de una carrera:

1) **Vocacionales:** "Eligen carreras y luego se dan cuenta de que no tienen la vocación necesaria".

2) **Económicas:** "A los estudiantes les es difícil financiar sus carreras, porque una parte importante de ellos proviene de los quintiles de menores recursos y un número

Universidad	Deserción al 1er año	¿Acreditada?
U. de los Andes	8,7%	5 años
UC del Maule	11,8%	5 años
U. del Bío-Bío	12,3%	5 años
U. de Talca	13,2%	5 años
U. Católica	13,7%	7 años
U. del Desarrollo	14,3%	5 años
U. de La Frontera	15,4%	5 años
U. Adolfo Ibáñez	15,6%	6 años
U. Diego Portales	15,6%	5 años
U. de Chile	16,1%	7 años
UC de la Sma. Concepción	16,7%	4 años
U. de Antofagasta	17,1%	4 años
U. Austral	17,3%	6 años
U. Adventista	17,8%	4 años
U. San Sebastián	17,8%	4 años
U. de Los Lagos	18,2%	3 años
U. de Santiago	18,4%	6 años
U. Mayor	18,6%	5 años
UC de Temuco	18,7%	3 años
UMCE	18,7%	4 años
U. de Concepción	18,8%	6 años
U. Finis Terrae	18,8%	5 años

- El estudio de aprobación presidencial hace inferencias a partir de una muestra, mientras que el estudio sobre deserción describe una población completa.

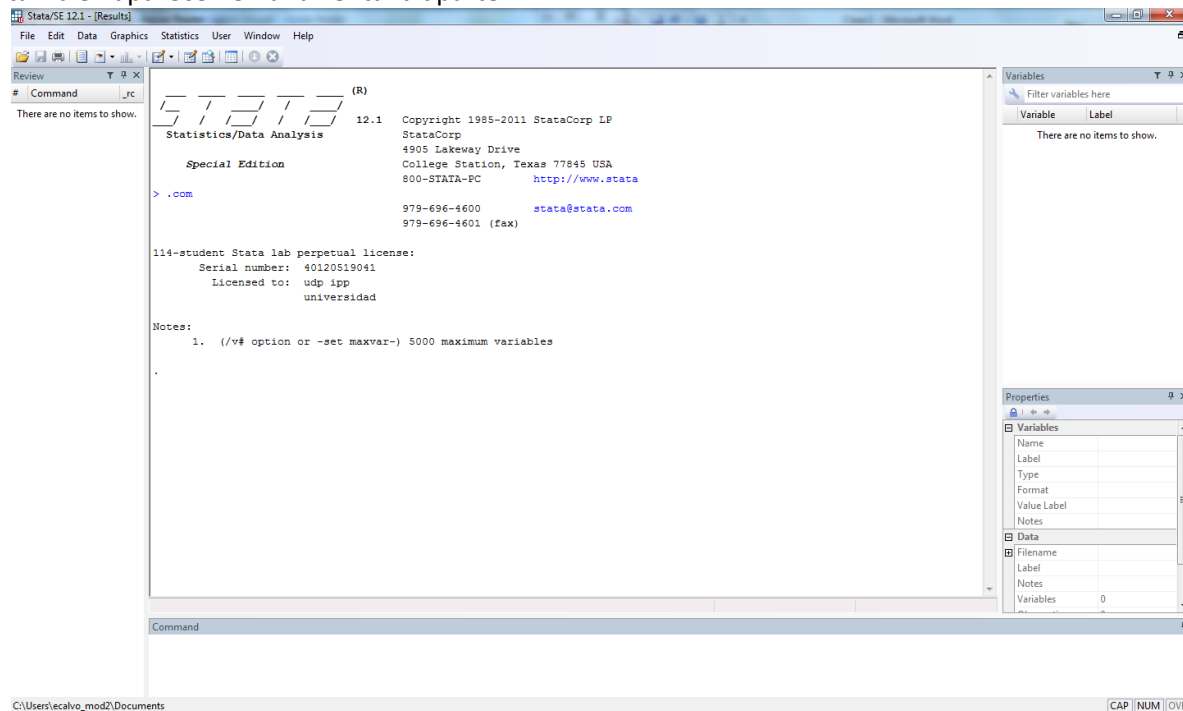
Ejemplo de Estadística Descriptiva e Inferencial Usando Casen 2011 en STATA

Stata es un programa computacional que permite administrar y analizar datos. El nombre viene de la combinación de las palabras *statistics* y *data*. Aprender a usarlo puede ser difícil en un principio, pero es una inversión que les abrirá oportunidades y permitirá ahorrar mucho tiempo en el futuro. Dado que seguirán usando Stata a lo largo de este curso, es muy recomendable instalar la última versión de Stata en sus computadores personales.

Al abrir Stata ves una pantalla con distintos paneles que puedes reubicar a tu gusto:

- Command (abajo): Permite ingresar los comandos.
- Results (centro): Registra todos los comandos ejecutados y resultados obtenidos. Las notas aparecen en azul, los errores en rojo y todo el resto en negro. No es posible borrar nada del panel central.
- Review (izquierda): Registra los comandos ingresados durante una sesión de Stata y permite ir construyendo un código automatizado. Los errores aparecen en rojo y pueden ser borrados.
- Variables (derecha, arriba): Describe las variables de la base de datos abierta.
- Properties (derecha, abajo): Describe la base de datos abierta.

Los íconos superiores con celdas o el menú Window permiten abrir el editor de datos en una nueva ventana. Los gráficos también aparecen en una ventana aparte.



Lo que haremos aquí es abrir un extracto de la base de datos CASEN 2011 y pedirle que resuma la variable edad. Deben descargar la “Base de Datos Inferencia Estadística” de la sección “Inferencia Estadística” del siguiente enlace:

http://www.esteban-calvo.com/es/difusion/_recursos/. Luego guarden la base de datos en el disco C o en un pendrive (traer pendrive para las próximas clases).

```
.use "C:\casen2011_ie.dta", clear
```

A continuación ocuparemos el comando <summarize>, que nos permite obtener una tabla resumen descriptivo de una variable, en este caso edad. La tabla que obtenemos muestra el número de observaciones que tiene la base de datos utilizada (4853), la media de la variable edad (43.7 años) y la respectiva desviación estándar (18.8). También muestra el valor mínimo y máximo de edad entre las observaciones (0 y 100).

```
.sum edad
```

Variable	Obs	Mean	Std. Dev.	Min	Max
edad	4853	43.64517	18.75686	0	100

Stata también nos permite realizar estadística inferencial. A modo de ejemplo aquí presentamos un <ttest> para comparar el promedio de ingreso (ytotaj) de hombres y mujeres. Lo que observamos es que existe una diferencia significativa entre los ingresos. La primera parte del test, hasta antes de *combined*, es estadística descriptiva, es decir, lo que ya han aprendido. Vemos que, claramente, los hombres reciben un ingreso más alto en esta muestra. Lo que aprenderemos en este curso, es lo que hay de *combined* hacia abajo. Lo que se hace en esa parte, es averiguar si es que esa diferencia es real o se debe al azar. Distinguir si algo es fruto del azar o no es justamente lo que hace la inferencia estadística. La última línea de la tabla nos dice que con un 99% de confianza podemos afirmar que la diferencia no se debe a la aleatoriedad.

```
. ttest ytotaj, by(sexo)
```

```
. ttest ytotaj, by(sexo)
```

Two-sample t test with equal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
hombre	2396	806409.8	26517.91	1298024	754409.3	858410.2
mujer	2457	454218.6	13742	681165.2	427271.5	481165.7
combined	4853	628100.8	15038.53	1047636	598618.4	657583.1
diff		352191.2	29654.49		294054.9	410327.4

diff = mean(hombre) - mean(mujer) t = 11.8765
 Ho: diff = 0 degrees of freedom = 4851

Ha: diff < 0 Ha: diff != 0 Ha: diff > 0
 Pr(T < t) = 1.0000 Pr(|T| > |t|) = 0.0000 Pr(T > t) = 0.0000

Aunque en este curso ocuparemos Stata de forma regular, el objetivo principal es aprender sobre inferencia estadística. Por lo tanto simplificaremos los ejercicios en Stata de modo que no tengan que hacer mucha administración de los

datos. En general trabajaremos con bases de datos con las variables ya creadas y sin datos perdidos. Sin embargo, esta es una buena oportunidad para desarrollar destrezas en el uso de Stata que vayan más allá de lo estrictamente necesario para realizar inferencias estadísticas. A lo largo de la carrera seguirán ocupando este software y muchos de ustedes lo seguirán ocupando durante su carrera.

Stata es un programa basado en comandos, mediante los cuales el usuario le dice lo que quiere hacer. La estructura típica de los comandos está compuesta de cuatro elementos: comando, nombre, condiciones y opciones. Para saber qué decirle a Stata, necesitamos responder cuatro preguntas que corresponden a los cuatro elementos mencionados:

- Comando: ¿Qué quiero hacer?
- Nombres: ¿Cuáles son los nombres de las variables o archivos que quiero ocupar?
- Condiciones: ¿Qué observaciones deseas ocupar?
- Opciones: ¿Tienes otros requisitos respecto a este programa?

Muchas veces es más fácil especificar con palabras normales lo que quiero hacer y luego intentar pedírselo a Stata. Para los usuarios avanzados, una de las ventajas de Stata es que permite a los usuarios crear sus propios comandos. Cada comando se debe ingresar en una línea separada, salvo que se esté trabajando en un archivo .do y se utilice /// al final de la línea para indicar que el comando sigue en la línea siguiente.

TAREA (5 MINUTOS)

Para seguir familiarizándose con Stata, la tarea que tienen para la siguiente clase es replicar las clases “Entering Data” y “Exploring Data”, disponibles en la siguiente página web de UCLA:

<http://www.ats.ucla.edu/stat/stata/notes/default.htm>. Está permitido colaborar con compañeros, pero cada uno tiene que entregar individualmente su tarea. La Tarea se entrega por email en formato Word. Deben cortar y pegar los contenidos desde Stata hacia Word. Para que no haya problemas de formato, utilizar márgenes estrechos, letra Courier New tamaño 10, interlineado simple y sin espacio entre párrafos. Si es necesario pueden cambiar la orientación de la página para que sea horizontal y no vertical. No se olviden de agregar su nombre a la tarea.

DISCUSIÓN DEL PROGRAMA DEL CURSO (15 MINUTOS)

El programa del curso describe los objetivos, contenidos, importancia, metodología, evaluación, bibliografía y políticas generales del curso. Es importante que lo lean en detalle y pregunten por aquello que no entiendan.

Es importante aclarar que se realizan tres clases a la semana: dos estilo cátedra y otro estilo taller en el laboratorio computacional. Intentamos que cada clase tenga una estructura dinámica y métodos de enseñanza variados. Como principio general, cada clase tiene tres secciones (introducción conceptual, desarrollo matemático y conclusión en Stata) y busca desarrollar dos tipos de aprendizaje (aprendizajes cognitivos en el ámbito del saber y aprendizajes procedimentales en el ámbito del saber hacer). La siguiente Tabla ilustra el diseño de una clase ideal, pero puede que no refleje el desarrollo de cada clase específica que se realice a lo largo del semestre.

Además de las clases existe una ayudantía por inscripción voluntaria que es importante que aprovechen, entre otras cosas, para repasar contenidos, resolver problemas y prepararse para las evaluaciones.

Secciones de la clase	Tipo de aprendizaje esperado	
	Cognitivo (saber)	Procedimental (saber hacer)
Introducción conceptual	<i>Indicador:</i> explicar conceptualmente para qué sirve y de qué se trata. <i>Método:</i> caso o ejercicio que ilustre para qué sirve y de qué se trata.	<i>Indicador:</i> analizar o aplicar conceptualmente. <i>Método:</i> ejercicio de análisis conceptual que ilustre qué puedo hacer.
Desarrollo matemático	<i>Indicador:</i> definir formalmente conceptos (fórmulas, estadísticos). <i>Método:</i> demostraciones matemáticas en pantalla.	<i>Indicador:</i> resolver problema matemático. <i>Método:</i> ejercicios prácticos con calculadora y papel.
Conclusión en Stata	<i>Indicador:</i> interpretar resultados. <i>Método:</i> demostraciones prácticas en Stata.	<i>Indicador:</i> obtener resultados. <i>Método:</i> ejercicios de aplicación práctica en Stata.

El material docente del curso Inferencia estadística incluye notas de clases disponibles en http://www.esteban-calvo.com/es/difusion/_recursos/. Si encuentran errores o tienen recomendaciones respecto a este material docente, por favor escribir a calvolab+ie@gmail.com.

Respecto a las evaluaciones, cabe destacar que las solemnes serán comunes a todas las secciones y que para los controles se calculará la nota de los cuatro mejores de un total de cinco, sin posibilidad de control recuperativo. Más detalles sobre las evaluaciones y otros aspectos del curso se pueden encontrar en el programa.