MODELOS DE PROBABILIDAD

Jorge Galbiati Riesco

EXPERIMENTOS ALEATORIOS

Considere las siguientes situaciones:

1. Se cuenta el número de naves que arriban a un puerto, por día.
2. Se le pregunta a un consumidor su marca preferida de leche.
3. Un fiscalizador examina declaraciones de impuesto y cuenta cuántas son erróneas.
4. Se observa el cambio mensual de un índice de precios.
5. Se toman los tiempos entre llegadas de clientes a la fila de una caja de banco.
6. Un investigador cuenta el número de partículas atómicas captadas por un instrumento.
7. Se extraen 20 peces de una lago, y se cuenta cuántos superan los 15 c. de largo
8. Se mide el peso de los contenidos de arroz en bolsas a la salida de una empacadora.
9. Un meteorólogo registra las temperaturas máximas diarias.
10. Un oftalmólogo registra el color de ojos de sus pacientes.
11. Un inspector registra el número de ítemes de frutas dañada en un cargamento.
12. Una persona compra un número de lotería y espera que sea el ganador.

Cada uno de estos ejemplos conesponde a un experimento aleatorio.

Un experimento aleatorio es un proceso que puede concretarse en al menos dos resultados posibles, con incertidumbre en cuanto a cuál de ellos tendrá lugar.

Una variable que es el resultado de un experimento, y que toma valores numéricos, es discreta si sólo puede tomar una cantidad numerable de valores.

En caso que tome valores en un intervalo de números reales, por lo tanto los valores no pueden enumerarse, se dice que la variable es contínua.

En los ejemplos, 2 y 10 no son variables numéricas. De las restantes, 1, 3, 6, 7, 11 y 12 son discretas. 6 y 12 tienen un rango extremadamente grande. 4, 5, 8 y 9 son contínuas, aunque las escalas en que se miden tienen limitaciones, por lo que lo que se registra resulta ser discreto.

ENFOQUE FRECUENTISTA DE LA PROBABILIDAD

A cada valor de una de estas variables se les pueden asignar **probabilidades,** que son una cuantificación de la certeza que se tiene de su ocurrencia. Son números, que convencionalmente toman valores entre 0 y 1, y además la suma total es 1.

Recordemos las tablas de frecuencia relativa, que mostraban la proporción de cada uno de los valores o clases de una variable, existentes en una muestra.

La muestra es la parte observada de una población, que es la que realmente interesa, aunque, por razones de recursos, no se observa completa.

Asociada a la población, también se podría construir una tabla de frecuencias relativas, si se conocieran todos los valores. Como generalmente no se conocen, sólo se puede postular una tabla de frecuencias relativas asociadas a la población.

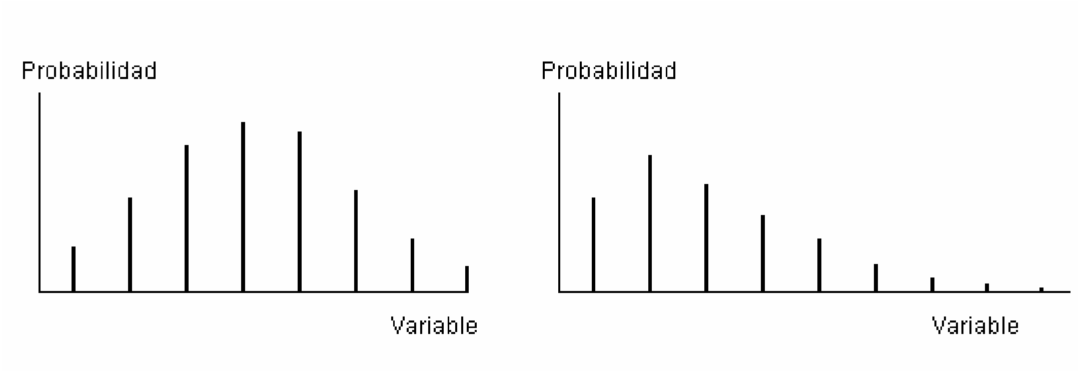
En tal caso, en lugar de tabla de frecuencias relativas, se denomina "función de probabilidad".

EJEMPLO 1. Se lanza un dado. Se observa la variable que representa el número resultante.
Si está balanceado entonces Prob(1) = P(2) = ... = P(6) = 1/6.

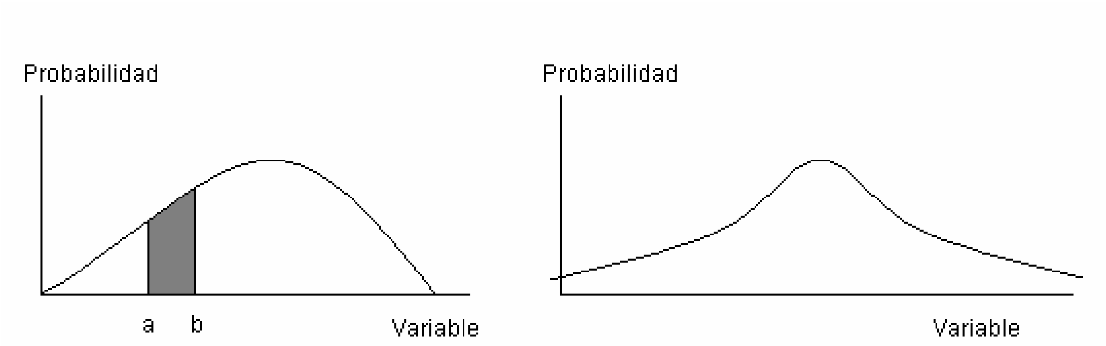
EJEMPLO 2: Se ha determinado que el número de errores por página del apunte de Estadística Básica es 0, 1 o 2, y sus respectivas probabilidades son

Prob(0) = 0.81 *Prob(1) = 0.17* *Prob(2) = 0.02*

Si la variable es discreta, cada valor tiene asociado una probabilidad, a nivel poblacional, equivalente a una frecuencia relativa, a nivel muestral. También se puede representar mediante un gráfico de barras, como se muestra en la figura siguiente.



Si la variable es continua, en el caso muestral se divide la tabla en intervalos, que pueden representarse mediante un histograma. En el caso poblacional, en cambio, se consideran intervalos infinitamente angostos, de modo que el perfil histograma toma la forma de una curva, como en la siguiente figura.



El **Valor esperado** de una variable es el promedio de todos sus valores, ponderados por sus respectivas probabilidades.

En el Ejemplo 2 el valor esperado de la variables es
(0)(0.81) + (1)(0.17) + (2)(0.02) = 0.21 errores por página.

EJEMPLO 3. Una ruleta contiene 38 casillas, numeradas 00, 0, 1, 2,.... 36. Una posibilidad consiste en apostar al suceso "el resultado es un número impar". Un jugador que realice esta apuesta ganará en 18 de los 38 resultados posibles. Puesto que hay 18 impares,

Prob (El resultado es un número par) = 18/38 = 9/19

Prob (El resultado es un número impar) = 18/38 = 9/19

Si un jugador realiza una apuesta \$1000, con lo cual recibe \$2000 si el resultado es un número impar y nada en caso contrario, podemos representar la "ganancia" (en pesos) del jugador como

\$1000 con probabilidad 9/19

- \$1000 con probabilidad (1-9/19) = 10/19

Entonces el valor esperado de su ganancia es

1000 x (9/19) + (-1000) x 10/19 = - 100 pesos. Espera perder \$100

Tanto en el caso discreto como continuo, existen modelos probabilísticos conocidos, que pueden aplicarse a determinados fenómenos, para representar, en forma aproximada, las proporciones de valores existentes en la población, es decir, sus probabilidades. Las funciones de probabilidad asociadas a estos modelos pueden escribirse como una ecuación matemática.

Es así como existen los modelos *Binomial*, *Geométrico*, *Hipergeométrico*, *Poisson*, *Exponencial*, *Normal*, *ji-cuadrado*, *t de Student*, *F*, *Gama*, *Beta*, entre muchísimos otros.

Por la forma en que se definieron las probabilidades, como frecuencias relativas, aplicadas a toda la población, resulta que son positivas y la suma de las probabilidades de todos los valores posibles de la variable que estamos describiendo, es igual a 1.

En el caso de los modelos de probabilidad contínuos, como la variable toma infinitos valores, la probabilidad asociada a un valor puntual es cero. Sólo pueden ser mayores que cero las probabilidades de intervalos de valores. Estas probabilidades pueden encontrarse, para algunos casos de uso común, en tablas, en las que se presentan habitualmente probabilidades de intervalos que parten desde menos infinito hasta una selección de valores de la variable.

Ejemplos:
Extraer al azar un número *n* de objetos, de una población en que hay de dos tipos (por ej, hombres y mujeres). El modelo ***binomial*** describe las probabilidades de obtener *0*, *1*, *2*, ..., o *n* de uno de los dos tipos.

El número de días de licencias médicas que se producen en una institución, o de fallas de un sistema computacional, en un mes, es aleatorio y podría representarse mediante un modelo ***Poisson***.

El número de declaraciones de impuesto que se debe revisar, hasta encontrar a un infractor, puede representarse probabilísticamente mediante un modelo ***geométrico***.

La probabilidad de encontrar un item defectuoso al inspeccionar un número determinado de ítemes de un lote pequeño, puede representarse mediante un modelo ***hipergeométrico***.

Una proporción de medidores de luz domiciliarios, que se encuentra descalibrados, en alguna región, podría representarse por un modelo ***beta***.

Los tiempos entre llegadas de clientes a una oficina de atención de público, pueden representarse por un modelo ***exponencial***.

El error respecto de una medida especificada, en un objeto producido por un proceso industrial, es una variable continua que puede representarse mediante un modelo ***normal***. También podría representar la dispersión, en torno a un valor promedio, de los puntajes de la prueba de selección universitaria.

En el caso de una muestra de valores de una variable podíamos calcular varios descriptores, que nos mostrarán algunas características numéricas de esa variable. De la misma forma, en el caso poblacional, también existirían estos descriptores, aunque, como no observamos toda la población, no podemos conocer sus valores. Estos valores, que son fijos pero desconocidos, se denominan ***parámetros*** de la población.

Entre ellos destacamos el promedio poblacional, que es el ***valor esperado*** de la variable, que ya definimos. Es una medida de centro, poblacional. También está la ***desviación estándar poblacional***, una medida de dispersión de la población.

Parte importante de la estadística consiste en la estimación de los parámetros de una población, a partir de evidencia muestral, y el estudio de las propiedades de la estimación. Por ejemplo, un valor esperado se puede estimar mediante un 0promedio

muestral; una desviación estándar poblacional, mediante una desviación estándar muestral.
Veremos algunos de estos modelos probabilísticos:

EL MODELO DE PROBABILIDAD BINOMIAL

Supongamos que un experimento aleatorio tiene sólo dos resultados posibles mutuamente excluyentes y conjuntamente exhaustivos, "éxito" y "fracaso", y que p es la probabilidad de obtener éxito en cada repetición. Si se realizan n repeticiones independientes, la distribución del número de éxitos, X, resultante se denomina **distribución binomial.** Su función de probabilidad es

$$P_x(x) = \frac{n!}{x!\,(n-x)!} p^x(1-p)^{n-x} \qquad \text{para } x= 0,1,2,\ldots,n$$

donde n! = 1 x 2 x 3 x 4 x...x n   y 0!=1

EJEMPLO 4:   Supongamos ahora que un agente de seguros tiene cinco contactos, y piensa que para cada uno la probabilidad de conseguir una venta es 0.4. La distribución del número de ventas, X es, entonces, binomial, con n = 5 y p = 0,4, es decir,

$$\text{Prob(X éxitos)} = \frac{5!}{X!\,(5-x)!}(0.4)^x(0.6)^{5-x} \quad \text{para } x = 0, 1,\ldots, 5$$

Las probabilidades para el número de éxitos (ventas logradas) son

$$P(0 \text{ éxitos}) = \frac{5!}{0!\,5!}(0,4)^0(0,6)^5 = (0,6)^5 = 0,078$$

$$P(1 \text{ éxitos}) = \frac{5!}{1!\,4!}(0,4)^1(0,6)^4 = (5)(0,4)(0,6)^4 = 0,259$$

$$P(2 \text{ éxitos}) = \frac{5!}{2!\,3!}(0,4)^2(0,6)^3 = (10)(0,4)^2(0,6)^3 = 0,346$$

$$P(3 \text{ éxitos}) = \frac{5!}{3!\,2!}(0,4)^3(0,6)^2 = (10)(0,4)^3(0,6)^2 = 0,230$$

$$P(4 \text{ éxitos}) = \frac{5!}{4!\,1!}(0,4)^4(0,6)^1 = (5)(0,4)^4(0,6) = 0,077$$

$$P(5 \text{ éxitos}) = \frac{5!}{5!\,0!}(0,4)^5(0,6)^0 = (0,4)^5 = 0,01$$

<u>EJEMPLO 5</u>**:** Una compañía recibe un gran cargamento de artículos, y decide aceptar el envío si en una muestra aleatoria de veinte artículos no hay más de uno defectuoso. Es decir, se acepta el cargamento si el número de artículos defectuosos es cero o uno, por lo que si Prob(X) es la función de probabilidad del número X de artículos defectuosos en la muestra, tenemos

P(aceptar el cargamento) = Prob(0) + Prob(1)

Supongamos que la proporción de artículos defectuosos en el cargamento es p = 0,1. Para n= 20, en la Tabla 1 del Apéndice, encontramos que las probabilidades de cero y un artículos defectuosos en la muestra son, respectivamente, Prob(0) = 0,1216 y Prob(1) = 0,2702. Por tanto, con esta regla de decisión, la probabilidad de que la compañía acepte él envío es

P(aceptar el cargamento) = 0,1216 + 0,2702 = 0,3918

Análogamente, si el 20% de los artículos del cargamento son defectuosos, es decir, si p=0,2, entonces,

P(aceptar el cargamento) = 0,0115 + 0,0576 = 0,0691

y para p= 0,3

P(aceptar el cargamento) = 0,0008 + 0,0068 = 0,0076

<u>EL MODELO DE PROBABILIDAD POISSON</u>

Supongamos que puede asumirse lo siguiente:

Para cada intervalo de tiempo muy pequeño de tiempo, la probabilidad de que ocurra un suceso en ese intervalo es aproximadamente proporcional a la amplitud del intervalo y no puede ocurrir dos o más sucesos en un intervalo.
Si lo anterior es cierto, puede probarse que la probabilidad de X ocurrencias en el intervalo de tiempo de 0 a T es

$$P(\text{x ocurrencias}) = \frac{e^{-\lambda}\lambda^{x}}{x!}$$

donde λ es el numero medio de ocurrencias entre 0 y T, y e = 2,71828 ... es la base de los logaritmos naturales. Este modelo probabilístico se denomina **Distribución de Poisson.**

EJEMPLO 6: Un estudio indica[9] que el número de huelgas anuales en una fábrica británica típica con 2.000 empleados, se puede representar por una distribución de Poisson con media λ = 0,4. La función de probabilidad del número de huelgas anuales X es, entonces,

$$\text{Prob (X huelgas)} = \frac{e^{-0.4}(0.4)^{x}}{X!} \qquad \text{para } x = 0, 1, 2, \ldots..$$

Podemos calcular ahora probabilidades para números concretos de huelgas anuales, usando (a partir de la Tabla 2 del Apéndice) $e^{-\lambda}$ = 0,6703.
La probabilidad de que no haya huelgas es

$$P(\text{0 huelgas}) = \frac{e^{-0.4}(0.4)^{0}}{0!} = \frac{(0.6703)(1)}{1} = 0.6703$$

Análogamente

$$P(\text{1 huelga}) = \frac{e^{-0.4}(0.4)^{1}}{1!} = \frac{(0.6703)(0.4)}{1} = 0.2681$$

$$P(\text{2 huelgas}) = \frac{e^{-0.4}(0.4)^{2}}{2!} = \frac{(0.6703)(0.16)}{21} = 0.0536$$

$$P(\text{3 huelgas}) = \frac{e^{-0.4}(0.4)^{3}}{3!} = \frac{(0.6703)(0.064)}{6} = 0.0071$$

$$P(\text{4 huelgas}) = e^{-0.4}(0.4)^{4} = (0.6703)(0.0256) = 0.0007$$

4! 24

Estas probabilidades pueden usarse para hallar la probabilidad de que el número de huelgas esté en un intervalo concreto. Por ejemplo, la probabilidad de que haya más de una huelga en un año es

P(más de 1 huelga) = 1 – P(0 huelgas) – P(1 huelga)

= 1 – 0.6703 – 0.2681 = 0.0616

<u>EJEMPLO 7</u>. La distribución de Poisson ha resultado ser muy útil en problemas de *líneas de espera o colas*. Los clientes llegan a una maquina fotocopiadora a una tasa media de dos cada cinco minutos En la práctica, se pueden representar los *procesos* de llegada de esta clase mediante una distribución de Poisson. Asumiendo que éste es el caso, representaremos por X el número de llegadas de clientes en un periodo de cinco minutos con lo cual X tiene una distribución de poisson con media $\lambda = 2$, y función de probabilidad

$$\text{Prob}(x) = \frac{e^{-2}2^x}{x!} \qquad \text{para } x = 0, 1, 2, \ldots$$

Las probabilidades para el número de llegadas en un período de cinco minutos son

$$P(0 \text{ llegadas}) = \frac{e^{-2}(2)^0}{0!} = \frac{(0.135335)(1)}{1} = 0.1353$$

$$P(1 \text{ llegadas}) = \frac{e^{-2}(2)^1}{1!} = \frac{(0.135335)(2)}{1} = 0.2707$$

$$P(2 \text{ llegadas}) = \frac{e^{-2}(2)^2}{2!} = \frac{(0.135335)(4)}{2} = 0.2707$$

---

y así sucesivamente. Así, por ejemplo, la probabilidad de que se produzcan más de dos llegadas en un periodo de cinco minutos es

P(X>2) = 1 – Prob(0) – Prob (1) – Prob (2) = 1 – 0.1353 – 0.2707 – 0.2707 = 0.3233

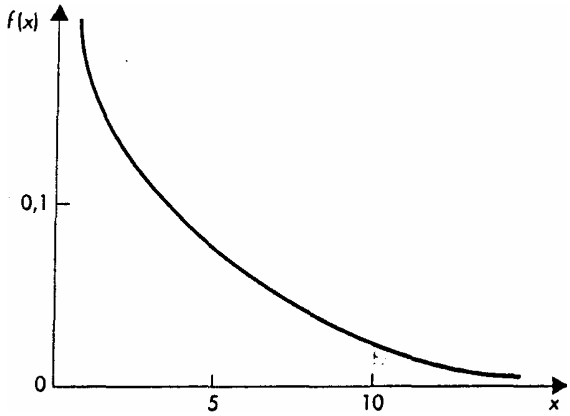
MODELO DE PROBABILIDAD EXPONENCIAL

Si la variable aleatoria *X* no puede tomar valores negativos y tiene función de densidad

$$f_X(x) = \frac{\exp\left(-x/\mu\right)}{\mu} \quad para \;\; x \geq 0$$

en que exp{} significa e elevado a lo que hay dentro del paréntesis, siendo *e* el número 2.71828...

y cero si x<0, donde μ es cualquier número positivo, entonces se dice que *X* sigue una **distribución exponencial.**



EJEMPLO 8. El tiempo de atención al cliente en un servicio de información de una biblioteca sigue una distribución exponencial, con un tiempo de servicio medio de cinco minutos. ¿Cuál es la probabilidad de que una consulta de un cliente dure más de diez minutos?

Sea *X* el tiempo de atención, en minutos. La función de densidad es

$$f_X(x) = \frac{\exp\left(-x/5\right)}{5} \quad para \;\; x \geq 0$$

La probabilidad que se pide es

P(X > 10) = 1 - *P(X* < 10) = 1 - $F_X(10)$ = 1 - (1 - $\exp\left(-10/5\right)$) = e^{-2} = 0,135335

La distribución exponencial está relacionada con la distribución de Poisson. Concretamente, si el número de ocurrencias de un suceso en un intervalo de tiempo sigue una distribución de Poisson con media λ, puede probarse que el tiempo que pasa entre dos ocurrencias consecutivas del suceso sigue una distribución exponencial de media $\mu = 1/\lambda$.

EL MODELO DE PROBABILIDAD NORMAL

La variación existe en todo fenómeno. Cuando tal variación se debe a una multitud de fuentes, que no son identificables, y que cada una aporta una pequeñísima contribución al fenómeno que estamos observando, suele ser apropiado el modelo normal para representar la variabilidad.

Su función de probabilidad está definida matemáticamente por la ecuación

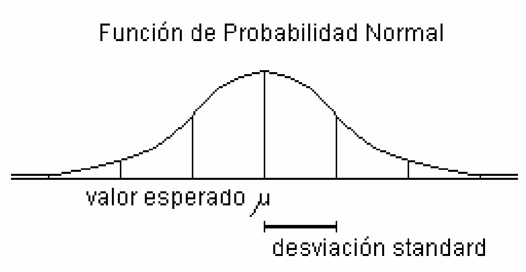
$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{ -\frac{1}{2} \frac{(x-\mu)^2}{\sigma^2} \right\}$$

en que exp{} significa e elevado a lo que hay dentro del paréntesis, siendo e el número 2.71828..., así como □ es el número 3.14159.... Los parámetros son □y es el valor esperado y□□ es la desviación estándar poblacional.

Un caso especial es el modelo probabilístico *normal estándar*, que tiene valor esperado cero y desviación estándar 1. La ecuación de la función de probabilidad es

$$p(x) = \frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{1}{2} x^2 \right\}$$

Es un modelo de probabilidad continuo, como se muestra en la figura siguiente:



Tiene la propiedad de que en el intervalo entre y se concentra una probabilidad de 0.683; en el intervalo entre y se concentra una probabilidad de 0.954; en el intervalo entre ☐ y se concentra una probabilidad de 0.997.

También presenta una propiedad muy útil, que dice que si X es una variable que sigue un modelo normal, con parámetros y , la variable transformada $\frac{X-\mu}{\sigma}$ sigue un modelo normal estándar. Esta propiedad seda en ambos sentidos. Es muy conveniente, por cuanto permite usar una tabla normal estándar para obtener probabilidades de cualquier otra normal.

Dos propiedades importantes relacionadas con el comportamiento de promedios (y proporciones) de muestras grandes. Como criterio, podemos considerar una muestra como "grande" cuando tiene a lo menos 25 elementos:

<u>Teorema del Límite Central</u>. Si se obtiene una muestra aleatoria de una población que obedece a cualquier modelo probabilístico, discreto o continuo, entonces si la muestra es suficientemente grande, su promedio sigue aproximadamente un modelo normal, con valor esperado igual al valor esperado de la población, y desviación standard igual a la desviación standard de la población, dividida por la raiz cuadrada del tamaño muestral.

Esta propiedad permite describir el comportamiento probabilístico de una media muestral con la normal, si se tiene una muestra grande, sin importar el comportamiento de la población de donde se obtuvieron los datos.

Ley de los Grandes Números. Si se obtiene una muestra aleatoria de una población que obedece a cualquier modelo probabilístico, discreto o continuo, entonces si la muestra es suficientemente grande, el promedio muestral se aproxima al promedio poblacional.

La ley de los grandes números nos indica que, si se toman muchas observaciones, a la larga los errores tienden a compensarse.

ESTIMACION POR INTERVALOS DE CONFIANZA

Se dieron dos ejemplos de estimación de parámetros. Debido a que estas consisten en un valor puntual, se denominan estimaciones puntuales. Obviamente no se espera que un valor estimado sea igual al verdadero parámetro que se pretende estimar. Sin embargo la estimación puntual no da cuenta de la magnitud del error que puede contener la estimación.

Por eso se prefiere hacer estimaciones mediante intervalos de confianza, que dan una estimación, una idea de la precisión de ésta, y un valor de probabilidad de que la estimación sea correcta, valor denominado *coeficiente de confianza* de la estimación.

Un intervalo de confianza tiene está determinado por los siguientes límites:
Estimador puntual $\pm$ (un factor de confianza)*(desviación estándar del estimador)

Por ejemplo, el porcentaje promedio del ingreso familiar gastado en viajes, de la población de familias, está entre 7.8% y 11.7%, con un coeficiente de confianza del 95%.

El coeficiente de confianza es un valor asociado a la probabilidad de que la estimación sea correcta, es decir, que el verdadero parámetro esté contenido en el intervalo. El ancho de intervalo es una medida de precisión: Mientras más angosto, más precisa la estimación.

Por ejemplo, si se tiene una muestra de observaciones, x_1, x_2,.., x_n , suficientemente grande (n mayor que 25, como criterio), un intervalo de confianza de coeficiente □□□□ en que □ es cercano a 100, está definido por los límites

$$\bar{x} - C_{\beta}\frac{s}{\sqrt{n}} \;\square\square\text{y}\; \bar{x} - C_{\beta}\frac{s}{\sqrt{n}}$$

en que $\bar{x}$ es el promedio muestral, *s* la desviación estándar muestral, n el tamaño de la muestra y C□ el factor asociado al coeficiente de confianza □, que se obtiene de una tabla normal. Esta fórmula se construye en base al teorema del límite central.

La siguiente tabla muestra algunos valores típicos del coeficiente de confianza y sus factores asociados:

Coeficiente de confianza □□□	Factor asociado C□
90	1.64
95	1.96
99	2.58

<u>EJEMPLO 9.</u> Se tomaron tres muestras distintas de la población de todas las comunas de Chile, de los datos del censo de 1992. Una muestra es de 30 observaciones, la segunda de 50 y la tercera de 70.

Se trata de una población con alta variabilidad, como se observa en las muestras (los valores poblacionales van de 131 habitantes, Antártica, hasta 334.305 habitantes, La Florida):

Obs.	muestra 1	muestra 2	muestra 3	Obs.	muestra 1	muestra 2	muestra 3
1	100739	22034	12372	36	22682	2877	
2	16418	22161	19807	37	1327	4589	
3	14214	25694	17129	38	35450	4569	
4	172637	195258	18740	38	13936	13324	
5	33307	27055	29846	40	47703	17289	
6	93377	22161	38527	41	5803	15560	
7	94393	342	44128	42	5322	5005	
8	3754	12347	7905	43	62191	6318	
9	48868	35450	48440	44	285212	13936	

10	15414	4589	1582	45	6686	4376
11	37628	131	39642	46	18425	35481
12	124833	6063	182715	47	30703	230770
13	14223	4072	47605	48	6346	42236
14	4416	11667	86357	49	13169	20786
15	22911	18338	23424	50	5803	46404
16	8879	1649	94393	51		86357
17	195258	28788	6686	52		18256
18	869	14163	13156	53		11129
19	10985	2011	12033	54		5559
20	10519	57049	27849	55		33307
21	10005	252176	17129	56		7905
22	20786	24685	1656	57		57294
23	28788	38518	2884	58		51788
24	69775	492	1654	59		10985
25	20191	14535	8841	60		10945
26	19400	84768	869	61		55604
27	14366	54403	80393	62		29719
28	17289	50532	30357	63		94393
29	131	13936	132750	64		20313
30	4841	25513	10599	65		25930
31		11667	50888	66		152907
32		46404	342	67		12309
33		12033	1199	68		50643
34		11527	14681	69		12321
35		492	38518	70		112653

Los siguientes cálculos corresponden a la primera muestra, con n=30 observaciones:
Promedio muestral = 40,973.8 hab.
Desviación estándar = 50,909.3 hab.
Desviación estándar dividido por raíz de 30 = 9294.7
calculamos un intervalo de confianza de coeficiente 95%. El factor es 1.96, según la tabla.

Entonces 1.96 veces 9294.7 es igual a 18,217.7
Los límites del intervalo son 40,973.8 - 18,217.7 = 22,756
y 40,973.8 + 18,217.7 = 59,192, redondeados al entero.
Es decir, con un 95% de confianza podemos afirmar que el promedio de toda la población está entre 22,756 y 59,192 habitantes.
Es un intervalo sumamente ancho, poco preciso, pero eso se debe a la alta variabilidad de esta variable.

La tabla siguiente contiene los intervalos de confianza de coeficientes 90%, 95% y 99%, calculados para las tres muestras. La unidad de medida de todos es "habitantes":

	Muestra		
Coef. confianza	No. 1, n=30 obs.	No. 2, n=50 obs.	No. 3, n=70 obs.
90%	25,730 a 56,217	20,949 a 47,589	26,961 a 44,151
95%	22,756 a 59,192	18,350 a 50,188	25,284 a 45, 828
99%	16,993 a 64,954	13,315 a 55,224	22,035 a 49,078

Se observa que a medida que aumenta la confianza, disminuye la presición. También se puede observar que con muestras más grandes, a igual coeficiente de confianza se obtienen intervalos más precisos.

A modo de comparación, el promedio de toda las comunas es 36,605 habitantes. Los promedios para las muestras de tamaños 30, 50 y 70 son, respectivamente, 40,974 ; 34,269 ; 35,556.

□

PREGUNTAS

1. ¿En qué situación se puede emplear el modelo normal para describir el comportamiento probabilístico de una población?

2. En cada caso dé un ejemplo de un experimento aleatorio asociado con el modelo probabilístico nombrado:

a) Binomial
b) Poisson

c) Hipergeométrico
d) Geométrico
e) Exponencial
f) Normal

3. ¿Qué diferencia hay entre un modelo de probabilidad discreto y un modelo de probabilidad contínuo?

4. ¿Què establece el Teorema del Límite Central?

5. ¿Qué expresa la ley de los grandes números?

6. Explique en qué consiste estimar un parámetro poblacionalmediante un intervalo de confianza de coeficiente C%.

7. Un consultorio estimó la duración promedio de las consultas de emergencia, mediante un intervalo de confianza de coeficiente 95%, utilizando una muestra de 35 casos. Se obtuvo el siguiente resultado: entre 8.75 y 22.86 minutos. Si se quiere volver a estimar, pero con mayor precisión, sin sacrificar el coeficiente de confianza, ¿Qué se debería hacer?