CAPÍTULO 8
INFERENCIA SOBRE UNA POBLACIÓN NORMAL

8.1 INTRODUCCIÓN

En el presente capítulo se introducen las técnicas básicas de inferencia sobre una población normal, sobre cuya media y desviación típica se desea obtener conclusiones a partir del análisis de los datos de una muestra extraída de la misma.

Para hacerlas más intuitivas, las cuestiones tratadas se introducen a partir del análisis de un problema concreto

Sobre dicho problema se recuerdan, en primer lugar, los conceptos de **hipótesis nula, y de riesgos de primera y segunda especie** que se vieron en el capítulo anterior y se introduce el concepto de **intervalo de confianza**, de gran importancia en la Inferencia Estadística.

Para abordar las cuestiones que se tratan en este tema, y en los siguientes, es necesario introducir previamente algunos conceptos básicos sobre la **distribución de las características muestrales**. En este capítulo se dan, de forma muy sintéticas, la ideas fundamentales al respecto, introduciéndose los resultados concretos a medida que se van necesitando posteriormente.

En este contexto se introduce también la distribución **t de Student**, que se manejará en todo el resto del texto.

Se expone también la forma de llevar a cabo los análisis mediante la utilización del paquete estadístico Statgraphics, ya manejado en anteriores capítulos.

8.2 UN EJEMPLO

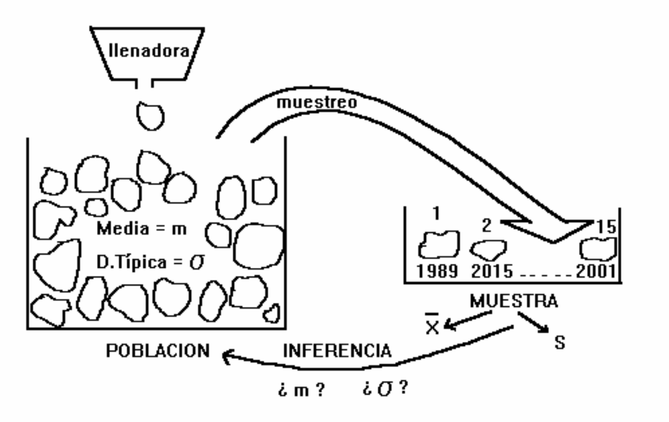
Una máquina llenadora de bolsas de malla de mandarinas, de las que se usan en los almacenes de confección de esta fruta, se regula para obtener un peso de 2000 gramos. Debido a una serie de causas de variabilidad (variación en el calibre de las frutas, imprecisión en las pesadas automáticas,...) es imposible obtener constantemente bolsas que pesen exactamente 2000 gramos; el peso obtenido es realmente una variable aleatoria, definida sobre la población de todas las bolsas que se confeccionan.

Se considera que la máquina está bien regulada si la media de dicha variable aleatoria es 2000. Para controlar si esto es así, se ha tomado al azar una muestra de 15 bolsas, cuyos pesos se recogen a continuación:

1989	2015	1962	2013	1983	1989	1992	2011
1958	2023	1980	1977	1994	2017	2001	

¿Qué puede decirse sobre la media y sobre la varianza de la variable estudiada?

En especial, ¿puede admitirse que la máquina está bien regulada o, por el contrario, hay evidencia de que la media es diferente de 2000 y debe, por tanto, procederse a reajustar la máquina?



8.3 NORMALIDAD DE LOS DATOS

La mayor parte de las técnicas de inferencia estadística clásicas sobre variables aleatorias continuas, asumen que las variables muestreadas se distribuyen normalmente. Antes, por tanto, de aplicar estas técnicas es aconsejable analizar hasta qué punto el modelo normal es adecuado, a la vista de los datos obtenidos, para analizar la pauta de variabilidad de la variable estudiada.

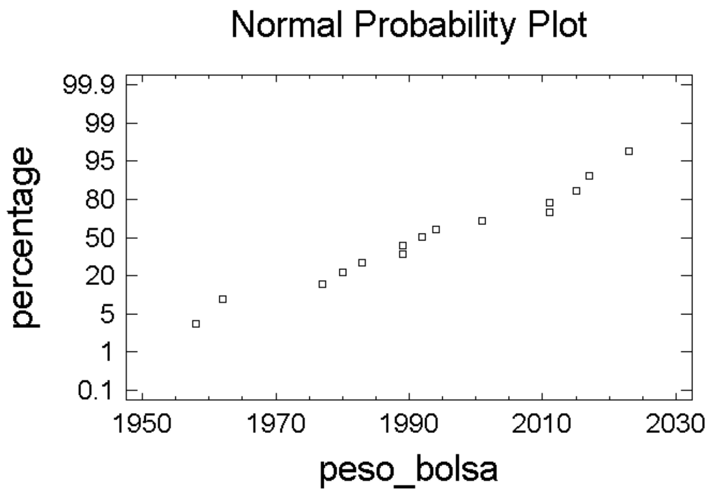
Nota importante: todo modelo matemático postulado para el estudio de un fenómeno real (sea la distribución normal, el péndulo matemático, o los rectángulos) implica necesariamente una simplificación y abstracción de ésta. (Carece por tanto de sentido la cuestión de si un determinado modelo es "verdadero" o "correcto" para aplicarlo a una realidad concreta! Sin embargo sí que es importante plantearse si la realidad estudiada es "lo suficientemente parecida" a la postulada por el modelo, como para que la aplicación de éste conduzca a resultados "válidos" o "útiles" en la práctica.

Autoevaluación: ¿Existen en el mundo muchos péndulos constituidos por un hilo inextensible y sin peso del que cuelga una masa puntual sin volumen que oscila sin rozamiento en el vacío?
¿Quiere ello decir que la fórmula del periodo del péndulo matemático no puede aplicarse nunca en la práctica? ¿Se podría aplicar dicha fórmula para calcular aproximadamente el periodo de un péndulo constituido por un trozo de papel colgado de una goma oscilando en medio de un vendaval?

Existen diversos tests estadísticos formales para estudiar la normalidad de unos datos (test Gi-dos, test de Kolmogorov, etcétera...). En nuestra opinión estos tests son de dudosa utilidad práctica, constituyendo "la respuesta correcta a una pregunta equivocada". En efecto, la pregunta a la que pretenden responder (¿proceden mis datos de una distribución normal teórica?) carece de sentido.

La pregunta realmente relevante sería: "¿es la pauta de variabilidad constatada en los datos lo suficientemente parecida a la postulada por el modelo matemático de la distribución normal, como para que pueda aplicarles los resultados teóricos deducidos para ésta?".

A esta pregunta, en vez de con test estadísticos formales de ajuste a distribuciones, se responde mejor a partir de una representación gráfica sencilla como un histograma o un gráfico en papel probabilístico normal. Estas representaciones pueden además detectar la presencia de datos anómalos (posiblemente la más grave y frecuente forma de no normalidad) y sugerir modelos alternativos en los casos en los que el normal aparezca como poco adecuado.



Dado que quince datos es un número muy reducido para elaborar un histograma, parece preferible en el ejemplo limitarse a obtener una representación de los datos en papel probabilístico normal. Tal como se ve en la figura, dicha representación no muestra ningún indicio claro de no normalidad, por lo que se asumirá el modelo normal como adecuado para el análisis posterior.

8.4 ANÁLISIS DESCRIPTIVO DE LA MUESTRA

La primera fase del análisis de cualquier conjunto de datos debe siempre consistir en un estudio descriptivo de los mismos, tanto mediante gráficos (histograma, diagrama Box-Whisker, Plot Normal), como calculando los principales "estadísticos (media, desviación típica, cuartiles, coeficientes de asimetría y curtosis, etc...)

Este análisis puede realizarse, por ejemplo, mediante Statgraphics que proporciona la salida que se recoge a continuación

A la vista de los resultados la hipótesis de normalidad sigue apareciendo como aceptable, puesto que los valores de los coeficientes de asimetría y curtosis estandarizados son bastante cercanos a cero.

Sin embargo, un simple análisis descriptivo no es suficiente para tomar decisiones a partir de los datos

Autoevaluación: la media muestral ha resultado igual a 1993.6, y es por tanto diferente de 2000 ¿Quiere ello decir que la máquina llenadora se ha desajustado y que, por tanto, hay que detener la producción y proceder a reajustar la máquina?
Procediendo de esta forma sistemáticamente:¿mejoraremos mucho nuestro proceso?, ¿lo empeoraremos?
Si la máquina llenadora está bien ajustada, o sea si su media es m = 2000: ¿Es seguro que en una muestra al azar de 15 bolsas se obtendrá siempre exactamente x =2000? ¿Es poco probable que pase eso? ¿Es prácticamente imposible que eso ocurra alguna vez?

Necesitamos por tanto algún procedimiento que nos permita obtener conclusiones sobre el valor de m en la población, a partir de la información que hemos obtenido en la muestra. Este tipo de procedimientos constituyen el objeto de la Inferencia Estadística.

8.5 ESTUDIO DE LA HIPÓTESIS M = 2000. ENFOQUE DEL PROBLEMA

Dado que la máquina se ajustó inicialmente a 2000 grs. y que no se ha producido ningún hecho excepcional en la planta (la muestra se ha tomado para hacer un control rutinario), parece razonable asumir como hipótesis de salida que m es igual a 2000 <u>a no ser que la información contenida en la muestra indique lo contrario</u>

A esta hipótesis que se toma como base de partida y que, en cierto sentido, refleja nuestro conocimiento previo de la situación, se le denomina en Inferencia Estadística **Hipótesis Nula H_0**

El proceso de razonamiento seguido para estudiar si esta hipótesis nula es o no admisible, es intuitivamente el siguiente:

Si m = 2000 $\Rightarrow$ $\bar{x}$ debería ser $\approx$ 2000

Por tanto, se admitirá que m=2000 si $\bar{x}$ es "cercana" a 2000
y se rechazará dicha hipótesis si $\bar{x}$ no es "cercana" a 2000

Pero...¿qué debería entenderse por "cercana"?

Para responder a esta pregunta es necesario cuantificar, de alguna manera, hasta qué punto un parámetro estimado a partir de una muestra, como $\bar{x}$, puede diferir por el azar del muestreo del valor m de ese parámetro en la población muestreada.

Seguidamente se aborda de forma muy sintética esta cuestión.

8.6 DISTRIBUCIONES EN EL MUESTREO

8.6.1 Introducción

Sea una población a cuyos individuos va asociada una variable aleatoria **X**. Para obtener conclusiones sobre la distribución de esta variable se obtiene una **muestra aleatoria** constituida por N individuos de la población. (Una muestra se dice que es aleatoria -en la terminología estadística se utiliza la expresión muestra aleatoria simple- si puede considerarse que todos los individuos de la población han tenido la misma probabilidad de estar incluidos en la muestra y si, además, dichos individuos han sido seleccionados independientemente unos de otros).

De una misma población es posible "a priori" obtener una gran cantidad de muestras diferentes de tamaño N (de hecho la cantidad es infinita si lo es la población original). Existe por tanto una **población de posibles muestras**, o sea una <u>nueva población cuyos individuos son dichas muestras</u>. A cada individuo de esta nueva población se le pueden hacer corresponder diferentes características numéricas (por ejemplo: la media muestral $\bar{x}$ o la desviación típica muestral s de la muestra considerada) que serán por tanto nuevas variables aleatorias.

Es muy importante comprender bien la idea expresada en el párrafo anterior. Esta idea implica que cualquier característica muestral (como $\bar{x}$ o s) es una nueva variable aleatoria, variable cuya distribución dependerá de la existente en la población muestreada y del tamaño de la muestra. Así, por ejemplo, la media muestral $\bar{x}$ es una variable aleatoria y, como tal, tiene una media y una desviación típica que dependerán, en general, de la distribución existente en la población de la que la muestra ha sido extraída.

Autoevaluación: ¿Tienen sentido las siguientes expresiones?
La media de la media muestral
La varianza de la media muestral
La media de la varianza muestral
La varianza de la varianza muestral
¿Podrías intuitivamente avanzar cuales crees que serán los valores de algunas de las características señaladas?

Cualquier función de los valores muestrales, como lo son por ejemplo la media muestral $\bar{x}$ o la desviación típica muestral s, se denomina un **estadístico**. De acuerdo con lo que acaba de exponerse todo estadístico es una variable aleatoria, cuya distribución dependerá en general de la distribución de la población y del tamaño de la muestra.

8.6.2 Distribución de $\bar{x}$

La media muestral se define, como sabemos, por la expresión:

$$\bar{x} = \frac{X_1 + \ldots + X_N}{N}$$

Si la población muestreada tiene de media m y de varianza σ^2, cada una de las X_i que constituye la muestra será el valor observado de una variable aleatoria con dichas media y varianza.

Como la media de una suma de variables es la suma de las medias, y la constante puede salir fuera al hallar esperanza matemática, se obtiene inmediatamente el siguiente importante resultado:

$$E\left(\bar{x}\right)=E\left(\frac{X_1+\ldots+X_N}{N}\right)=\frac{E(X_1)+\ldots+E(X_N)}{N}=\frac{m+\ldots+m}{N}=m$$

Por tanto, **la media de la media muestral es la media poblacional**.

Por otra parte, aplicando las dos propiedades de la varianza vistas en el apartado 5.7.1, se obtiene:

$$\mathrm{Var}\left(\bar{x}\right)=\mathrm{Var}\left(\frac{X_1+\ldots+X_N}{N}\right)=\frac{\mathrm{Var}(X_1)+\ldots+\mathrm{Var}(X_N)}{N^2}=\frac{\sigma^2+\ldots+\sigma^2}{N^2}=\frac{N\sigma^2}{N^2}=\frac{\sigma^2}{N}$$

Por tanto, **la varianza de la media muestral es la varianza de la población dividida por el tamaño N de la muestra**.

En consecuencia a medida que aumenta N disminuirá $\sigma^2(\bar{x})$ y resultará más improbable que $\bar{x}$ tome valores que difieran mucho de la media poblacional m. Este resultado justifica la idea intuitiva habitual de considerar más "fiables" las conclusiones obtenidas a partir de una muestra grande que las derivadas de una muestra pequeña.

Digamos, por último, que como una consecuencia inmediata del Teorema Central del Límite visto en el capítulo anterior, la media muestral $\bar{x}$, al ser el resultado de sumar una serie de variables independientes, tiende a distribuirse normalmente, aunque la población de la que se haya extraído la muestra no siga una distribución normal.

Nota:en el caso particular, muy importante en la práctica, de que en la población muestreada la variable se distribuya normalmente, $\bar{x}$ se distribuye normalmente, sea cualsea el tamaño N de la muestra, e, en consecuencia, $\dfrac{\bar{x}-m}{\sigma/\sqrt{N}}$ sigue una distribución normal tipificada.

*Evaluación: Se extrae una muestra de tamaño **N** de una población normal de desviación típica σ=3. Calcular cuanto debe valer como mínimo **N** si se desea tener una probabilidad inferior al 1% de que la diferencia entre la media muestral y la media poblacional sea superior a 1 (Ver respuesta al final del tema)*

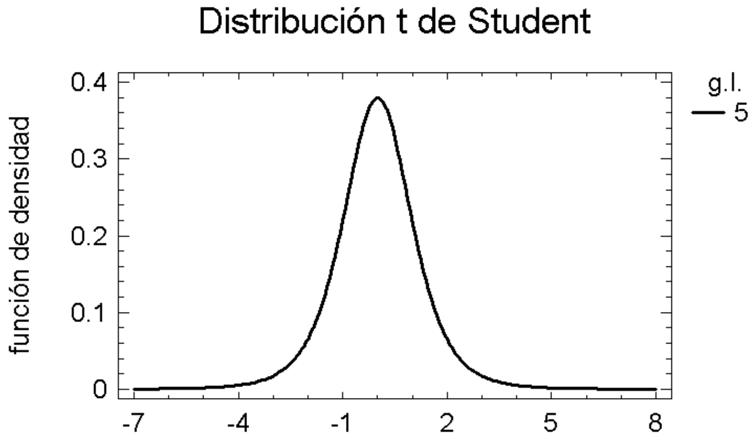
8.6.3 La distribución t de Student

En la inferencia respecto a medias en poblaciones normales desempeña un papel fundamental la distribución t de Student. (El nombre procede del seudónimo con que firmaba sus trabajos el matemático Gosset, que la desarrolló en el curso de sus trabajos centrados en la industria cervecera).

Siendo **X** una variable N[0,1] e **Y** una variable Gi-dos con ν grados de libertad independiente de la anterior, se dice que **t** sigue una distribución t de Student con ν grados de libertad si:

$$t = \frac{X}{\sqrt{Y/\nu}}$$

En la figura siguiente se refleja la forma de la función de densidad de una variable t de Student con 5 grados de libertad.



Como puede apreciarse la distribución es simétrica respecto al origen, que es la media de la variable, recordando mucho a una N[0,1].

En realidad la varianza de la t de Student es algo mayor que 1 (concretamente $\sigma^2 = \nu/(\nu-2)$, asumiendo que $\nu>2$), pero tiende a 1 a medida que aumentan los grados de libertad. De hecho cuando ν tiende a 4 la variable t de Student tiende a distribuirse como una N[0,1].

La tabla de la siguiente hoja refleja, para diferentes valores de α y ν, los valores t_α tales que la probabilidad de que una variable t de Student con ν grados de libertad sea superior a t_α es igual a α.

Autoevaluación: Obtener un valor x tal que la probabilidad de que una t de Student con 10 grados de libertad sea en valor absoluto mayor que x, sea igual al 5%.

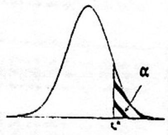
La importancia de la distribución t de Student en la Inferencia Estadística radica en el siguiente resultado: siendo $\overline{x}$ y s la media y la desviación típica de una muestra de tamaño N extraída de una población normal de media m, el estadístico:

$$t = \frac{\overline{x} - m}{s/\sqrt{N}}$$ sigue una distribución t de Student con N-1 grados de libertad.

Nota: Apréciese la analogía entre la expresión anterior y la vista al final de 8.6.2. El hecho de sustituir en aquélla una desviación típica poblacional σ (generalmente desconocida) por su estimación muestral s, se traduce en que el estadístico pasa de tener una distribución N[0,1] a una t de Student con N-1 grados de libertad (que son los grados de libertad con que está estimada σ).

D. E. I. O.

DISTRIBUCIÓN t



Grados de libertad - n	Probabilidad de una cola - α												
	0.0005	0.001	0.005	0.010	0.025	0.050	0.100	0.200	0.250	0.3	0.400	0.450	0.475
1	636.578	318.289	63.656	31.821	12.706	6.314	3.078	1.376	1.000	0.727	0.325	0.158	0.079
2	31.600	22.328	9.925	6.965	4.303	2.920	1.886	1.061	0.816	0.617	0.289	0.142	0.071
3	12.924	10.214	5.841	4.541	3.182	2.353	1.638	0.978	0.765	0.584	0.277	0.137	0.068
4	8.610	7.173	4.604	3.747	2.776	2.132	1.533	0.941	0.741	0.569	0.271	0.134	0.067
5	6.869	5.894	4.032	3.365	2.571	2.015	1.476	0.920	0.727	0.559	0.267	0.132	0.066
6	5.959	5.208	3.707	3.143	2.447	1.943	1.440	0.906	0.718	0.553	0.265	0.131	0.065
7	5.408	4.785	3.499	2.998	2.365	1.895	1.415	0.896	0.711	0.549	0.263	0.130	0.065
8	5.041	4.501	3.355	2.896	2.306	1.860	1.397	0.889	0.706	0.546	0.262	0.130	0.065
9	4.781	4.297	3.250	2.821	2.262	1.833	1.383	0.883	0.703	0.543	0.261	0.129	0.064
10	4.587	4.144	3.169	2.764	2.228	1.812	1.372	0.879	0.700	0.542	0.260	0.129	0.064
11	4.437	4.025	3.106	2.718	2.201	1.796	1.363	0.876	0.697	0.540	0.260	0.129	0.064
12	4.318	3.930	3.055	2.681	2.179	1.782	1.356	0.873	0.695	0.539	0.259	0.128	0.064
13	4.221	3.852	3.012	2.650	2.160	1.771	1.350	0.870	0.694	0.538	0.259	0.128	0.064
14	4.140	3.787	2.977	2.624	2.145	1.761	1.345	0.868	0.692	0.537	0.258	0.128	0.064
15	4.073	3.733	2.947	2.602	2.131	1.753	1.341	0.866	0.691	0.536	0.258	0.128	0.064
16	4.015	3.686	2.921	2.583	2.120	1.746	1.337	0.865	0.690	0.535	0.258	0.128	0.064
17	3.965	3.646	2.898	2.567	2.110	1.740	1.333	0.863	0.689	0.534	0.257	0.128	0.064
18	3.922	3.610	2.878	2.552	2.101	1.734	1.330	0.862	0.688	0.534	0.257	0.127	0.064
19	3.883	3.579	2.861	2.539	2.093	1.729	1.328	0.861	0.688	0.533	0.257	0.127	0.064
20	3.850	3.552	2.845	2.528	2.086	1.725	1.325	0.860	0.687	0.533	0.257	0.127	0.063
21	3.819	3.527	2.831	2.518	2.080	1.721	1.323	0.859	0.686	0.532	0.257	0.127	0.063
22	3.792	3.505	2.819	2.508	2.074	1.717	1.321	0.858	0.686	0.532	0.256	0.127	0.063
23	3.768	3.485	2.807	2.500	2.069	1.714	1.319	0.858	0.685	0.532	0.256	0.127	0.063
24	3.745	3.467	2.797	2.492	2.064	1.711	1.318	0.857	0.685	0.531	0.256	0.127	0.063
25	3.725	3.450	2.787	2.485	2.060	1.708	1.316	0.856	0.684	0.531	0.256	0.127	0.063
26	3.707	3.435	2.779	2.479	2.056	1.706	1.315	0.856	0.684	0.531	0.256	0.127	0.063
27	3.689	3.421	2.771	2.473	2.052	1.703	1.314	0.855	0.684	0.531	0.256	0.127	0.063
28	3.674	3.408	2.763	2.467	2.048	1.701	1.313	0.855	0.683	0.530	0.256	0.127	0.063
29	3.660	3.396	2.756	2.462	2.045	1.699	1.311	0.854	0.683	0.530	0.256	0.127	0.063
30	3.646	3.385	2.750	2.457	2.042	1.697	1.310	0.854	0.683	0.530	0.256	0.127	0.063
40	3.551	3.307	2.704	2.423	2.021	1.684	1.303	0.851	0.681	0.529	0.255	0.126	0.063
60	3.460	3.232	2.660	2.390	2.000	1.671	1.296	0.848	0.679	0.527	0.254	0.126	0.063
120	3.373	3.160	2.617	2.358	1.980	1.658	1.289	0.845	0.677	0.526	0.254	0.126	0.063
∞	3.290	3.090	2.576	2.326	1.960	1.645	1.282	0.842	0.674	0.524	0.253	0.126	0.063
	0.001	0.002	0.010	0.020	0.050	0.100	0.200	0.400	0.500	0.6	0.800	0.900	0.950
	Probabilidad de dos colas - 2 α												

8.7 CONTRASTE DE LA HIPÓTESIS m = 2000

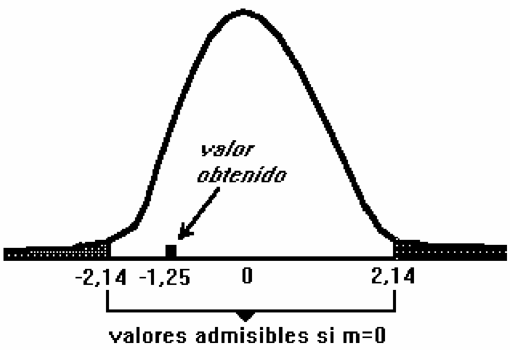
Al final del Apartado 8.5 preguntábamos hasta qu´re punto debía diferir lamedia muestral $\overline{x}$ del valor hipotético 2000 para que fuera razonable rechazar la Hipótesis Nula H_0 m = 2000

La solución consiste en basarse en el estadístico $\mathbf{t_{calc}} = \dfrac{\overline{x} - 2000}{s/\sqrt{N}}$ (donde N=15 es el tamaño de la muestra) que, de acuerdo con lo visto al final del apartado 8.6.3, seguirá una distribución t de Student con 14 grados de libertad (y será por tanto próximo a 0) si es cierta la hipótesis de que m es igual a 2000.

Por otra parte, si m es diferente de 2000, la t_{calc} tomará en general valores más lejanos de cero (positivos si m>2000 ó negativos si m<2000) que los esperables para una variable t de Student. Por tanto, una forma razonable de proceder será **rechazar la hipótesis nula si la t_{calc} resulta <u>demasiado grande</u> (en valor absoluto) para ser una t de Student**

Y ¿cómo podemos cuantificar hasta qué punto la t_{calc} es "demasiado grande" para ser una t de Student? $\Rightarrow$ calculando la probabilidad de que una t de Student tome valores tan grandes o más que la t_{calc}. A esta probabilidad, que puede obtenerse en las tablas de la t, se le denomina el **p-value**, y contra más baja sea mayor será la evidencia respecto a la falsedad de H_0.

Frecuentemente se opera en Estadística rechazando la hipótesis nula si el p-value es inferior a 0.05. Ello es equivalente, en nuestro ejemplo, a rechazar la hipótesis nula si el estadístico t resulta mayor (en valor absoluto) al valor correspondiente al punto $t_{14}(0.05)$ que, como puede consultarse en la tabla correspondiente, es 2.14



Como en el ejemplo:

$$\left|t_{calc}\right| = \left|\frac{\overline{x} - 2000}{s/\sqrt{N}}\right| = \left|\frac{1993.6 - 2000}{19.8/\sqrt{15}}\right| = \left|-1.25\right| < 2.14$$

se deduce que la hipótesis nula m = 2000 es aceptable, o sea, que es compatible con los datos observados en la muestra.

Autoevaluación: contra más pequeño sea el p-value, más significativa se dice que es la diferencia entre m y 2000. Que una diferencia sea muy significativa estadísticamente, ¿quiere decir que es muy grande? Cuando en el ejemplo considerado se admite la hipótesis nula, ¿quiere decir que hemos demostrado estadísticamente que m es igual a 2000? (Ver respuesta en el Anejo al final del Tema)

En síntesis, la operativa general propuesta para analizar si es admisible la hipótesis H_0: m = m_0 es la siguiente:

Calcular $t_{calc} = \dfrac{\overline{x} - m_0}{s/\sqrt{N}}$ y rechazar H_0 si $|t_{calc}| > t_{N-1}(\alpha)$, donde $t_{N-1}(\alpha)$ es un valor, que se busca en tablas tal que $P(|t_{N-1}| > t_{n-1}(\alpha)) = \alpha$, donde en nuestro ejemplo hemos operado con α = 0.05

Riesgos de 1ª y 2ª especie

α será por lo tanto la probabilidad que tenemos, operando de esta forma, de equivocarnos rechazando H_0 cuando es cierta. A α se le denomina, como sabemos, Riesgo de 1ª especie.

(Nota: en general en Estadística la palabra "riesgo" es un término técnico que significa "probabilidad de cometer un error")

En general se utilizan valores de α=0.05 (como en el ejemplo) ó α=0.01 (si se requiere tener un riesgo de 1ª especie muy bajo).

Autoevaluación: ¿No sería mejor operar con riesgos de 1ª especie lo más bajos posibles? ¿Qué problema se presentaría en la práctica si queremos reducir al máximo la probabilidad de equivocarnos rechazando la H_0 cuando sea cierta?

Si se opera con α muy bajo (en el ejemplo: si sólo queremos reajustar la máquina si estamos muy seguros de que se ha desajustado), se aumenta a cambio el Riesgo de 2ª especie, que es, como sabemos, la probabilidad de aceptar H_0 cuando es falsa (no reajustar la máquina pese a haberse desajustado).

Se considera generalmente que α=0.05 es un compromiso razonable entre ambos riesgos.

Nota técnica importante: el ingeniero no debe, sin embargo, obsesionarse ciegamente en sus estudios con unos valores concretos de α (0.05 ó 0.01) La utilización de unos valores límites o "críticos" para el p-value que separan los resultados "significativos" de los "no significativos" no es más que un anacronismo, reflejo de épocas en las que el cálculo exacto de estos p-values se hallaba fuera del alcance del investigador, que sólo podía hacerse una idea al respecto comparando los valores por él obtenidos con los reflejados en unas tablas que generalmente se limitaban a estos dos niveles. Es el p-value, por tanto, lo que refleja el grado de evidencia de unos resultados contra la H_0 y, en consecuencia, lo que debería acompañar al análisis de dichos resultados, y no sólo la constatación de si resulta superior o inferior al 5%. Porque ¿qué diferencia hay, en la práctica, entre un p-value del 4.9% o del 5.1%?

8.8 INTERVALO DE CONFIANZA PARA m

Que la H_o m=2000 sea admisible para los datos obtenidos, no quiere decir que hayamos demostrado que m es igual a 2000. De hecho, muchos otros valores hipotéticos de m también hubieran resultado admisibles con los datos obtenidos.

¿Cómo podemos, a partir de la muestra, construir un intervalo que tenga a priori una probabilidad elevada (1-α) de contener el valor desconocido m de la media poblacional?

Como $\frac{\bar{x}-m}{s/\sqrt{N}}$ se distribuye como una t_{N-1}, y siendo $P(|t_{N-1}|>t_{n-1}(\alpha))=\alpha$, se cumplirá que:

$$P\left(-t_{N-1}(\alpha)<\frac{\bar{x}-m}{s/\sqrt{N}}<+t_{N-1}(\alpha)\right)=1-\alpha$$

de donde se obtiene inmediatamente

$$P\left(\bar{x} - t_{N-1}(\alpha)\frac{s}{\sqrt{N}} < m < \bar{x} + t_{N-1}(\alpha)\frac{s}{\sqrt{N}}\right) = 1-\alpha$$

Por tanto $\left[\bar{x} - t_{N-1}(\alpha)\frac{s}{\sqrt{N}}\,;\, \bar{x} + t_{N-1}(\alpha)\frac{s}{\sqrt{N}}\right]$ constituye un intervalo de confianza con una probabilidad 1-α de contener a m. (A 1-α se le denomina también "nivel de confianza").

Autoevaluación: ¿qué interpretación práctica tiene la probabilidad 1-α asociada a un determinado intervalo de confianza? (Ver respuesta en el Anejo al final del Tema)

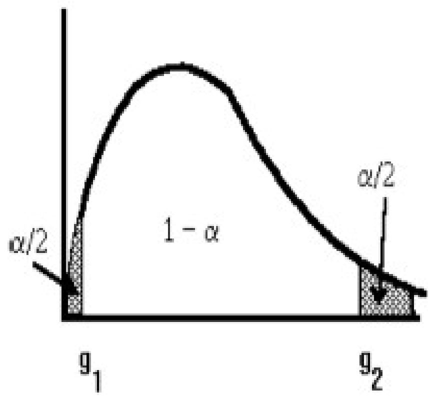
En el ejemplo: $\left[1993.6 - 2.14\frac{19.8}{\sqrt{15}}\,;1993.6 + 2.14\frac{19.8}{\sqrt{15}}\right] = [1982.7\,;2004.5]$ constituye un intervalo de confianza para la media m del proceso, para un nivel de confianza del 95%.

Nota: como es lógico, el intervalo contiene el valor m=2000, lo que era de esperar dado que la hipótesis m=2000 ha resultado aceptable.

8.9 INTERVALO DE CONFIANZA PARA σ

Siendo s^2 la varianza muestral obtenida en una muestra de tamaño N extraída de una población Normal de varianza σ^2, se demuestra que $(N-1)\frac{s^2}{\sigma^2}$ sigue una distribución Gi-dos con N-1 grados de libertad.(La distribución Gi-dos fue estudiada en el capítulo anterior)

A partir de este resultado es sencillo obtener un intervalo de confianza para σ, tal como se expone a continuación.



En efecto, como $(N-1)\frac{s^2}{\sigma^2}$ sigue una distribución χ^2 con (N-1) grados de libertad, de la tabla de la χ^2 es posible obtener dos valores, g_1 y g_2, tales que

$$P(g_1 < \chi^2_{N-1} < g_2) = 1-\alpha$$

(Por ejemplo, $P(5.63 < \chi^2_{14} < 26.1) = 0.95$)

Por tanto, $P(g_1 < (N-1)\frac{s^2}{\sigma^2} < g_2) = 1-\alpha$

de donde se obtiene

$$P\left((N-1)\frac{s^2}{g_2} < \sigma^2 < (N-1)\frac{s^2}{g_1}\right) = 1-\alpha$$

$\left[(N-1)\frac{s^2}{g_2} ; (N-1)\frac{s^2}{g_1} \right]$ constituirá, en consecuencia un intervalo de confianza para la varianza poblacional σ^2, y $\left[\sqrt{(N-1)\frac{s^2}{g_2}} ; \sqrt{(N-1)\frac{s^2}{g_1}} \right]$ será un intervalo de confianza para la desviación típica poblacional σ.

En el ejemplo $\left[\sqrt{(15-1)\frac{19.8^2}{26.1}} ; \sqrt{(15-1)\frac{19.8^2}{5.63}} \right]$ = [14.5 ; 31.2] será el intervalo de confianza, con un nivel de confianza del 95%, para σ.

8.10 ANÁLISIS MEDIANTE STATGRAPHICS

La obtención de intervalos de confianza para la media y la varianza de una población, así como el contraste de una hipótesis sobre el valor de la media, puede llevarse a cabo mediante la opción *describe ... numeric data ... one-variable analysis* del paquete Statgraphics.

La única entrada que hay que dar es el nombre de la variable en la que están guardados los valores muestrales. (En el ejemplo a dicha variable, que contiene los 15 valores observados para los pesos de las bolsas, se le ha denominado *peso-bolsas*)

El programa proporciona en primer lugar los estadísticos básicos descriptivos de la muestra. Mediante el botón derecho del ratón es posible solicitar un test de hipótesis para un determinado valor m_0 de la media y el riesgo α de 1ª especie que se desee (m_0 y α se fijan dentro de *pane_options*).

El resultado es:

También mediante el botón derecho del ratón es posible solicitar intervalos de confianza, con un determinado nivel de confianza, para m y para σ:

8.11 INTERVALO DE CONFIANZA PARA UNA PROPORCIÓN

Vamos a estudiar la forma de obtener un Intervalo de Confianza para la proporción **p** de individuos de una población que verifican cierto suceso, a partir de la frecuencia X observada en una muestra de tamaño N

Sabemos que, si N no es pequeña, la variable Binomial X se distribuye de forma aproximada como una Normal de media Np y $\sigma = \sqrt{Np(1-p)}$

En consecuencia, la frecuencia relativa con la que se ha observado el suceso, p*=X/N, se distribuirá aproximadamente como una Normal de media p y σ igual a $\sqrt{\frac{p(1-p)}{N}}$

Por lo tanto, si $z_{\alpha/2}$, obtenido de la tabla de la Normal, es el valor que verifica:
Prob(- $z_{\alpha/2}$ < N(0,1) < $z_{\alpha/2}$) = 1- α, se cumplirá que:

$$\mathrm{Prob}\left(-z_{\alpha/2} < \frac{P^*-p}{\sqrt{p(1-p)/N}} < z_{\alpha/2}\right) = \mathrm{Prob}\left(-z_{\alpha/2}\sqrt{\frac{p(1-p)}{N}} < P^*-p < z_{\alpha/2}\sqrt{\frac{p(1-p)}{N}}\right) = \mathrm{Prob}\left(P^*-z_{\alpha/2}\sqrt{\frac{p(1-p)}{N}} < p < P^*+z_{\alpha/2}\sqrt{\frac{p(1-p)}{N}}\right) = 1-\alpha$$

Con lo que los dos límites de la desigualdad de la derecha definirían un Intervalo de Confianza para p con un nivel de confianza 1- α.

Como en la expresión del intervalo de confianza aparece el valor de p, que es desconocido, se sustituye por su estimación p*, siendo también conveniente sustituir el valor $z_{\alpha/2}$, asociado a una Normal, por el percentil correspondiente $t_{n-1,\alpha/2}$ de una t de Student con N-1 grados de libertad. La expresión del Intervalo de Confianza para p queda, por tanto:

$$\left(P^* - t_{n-1}^{\alpha/2}\sqrt{\frac{P^*(1-P^*)}{N}} \;\; ; \;\; P^* + t_{n-1}^{\alpha/2}\sqrt{\frac{P^*(1-P^*)}{N}}\right)$$

Evaluación: se desea hacer una encuesta para estimar el % de votantes a los diferentes partidos en unas próximas elecciones. ¿Qué tamaño mínimo N debe tener la muestra si se desea que la probabilidad de cometer un error superior a 5 puntos porcentuales en la estimación de una determinada proporción P sea inferior al 1%? (Ver respuesta al final del capítulo)

8.A AUTOEVALUACIONES RESUELTAS Y EJERCICIOS

8.A.1 Respuesta a algunas Autoevaluaciones

*Evaluación: Se extrae una muestra de tamaño **N** de una población normal de desviación típica σ=3. Calcular cuanto debe valer como mínimo **N** si se desea tener una probabilidad inferior al 1% de que la diferencia entre la media muestral y la media poblacional sea superior a 1 (Ver respuesta al final del tema)*

Como $\frac{\bar{x}-m}{\sigma/\sqrt{N}} = \sqrt{N}\frac{\bar{x}-m}{3} \sim N(0,1) \Rightarrow$

$P(|\bar{x}-m|>1) = P\left(\left|\sqrt{N}\frac{\bar{x}-m}{3}\right| > \sqrt{N}/3\right) = P\left(|N(0,1)| > \sqrt{N}/3\right)$ ha de ser < 0.01

En la tabla de la normal se ve que para que se cumpla esto $\sqrt{N}/3$ ha de ser > 2.575 $\Rightarrow$

$$N > (3x2.575)^2 \Rightarrow N = 60$$

Autoevaluación: contra más pequeño sea el p-value, más significativa se dice que es la diferencia entre m y 2000. Que una diferencia sea muy significativa estadísticamente, ¿quiere decir que es muy grande? Cuando en el ejemplo considerado se admite la hipótesis nula, ¿quiere decir que hemos demostrado estadísticamente que m es igual a 2000?

Un error muy frecuente, es la confusión entre significación estadística e importancia práctica. El problema es de naturaleza semántica, y deriva de utilizar, para designar un concepto técnico estadístico concreto, un vocablo -"significativo"- que tiene un sentido diferente en el lenguaje habitual. Si la diferencia entre m y 2000 es "muy significativa estadísticamente", la interpretación práctica correcta es que resulta casi seguro que dicha diferencia no es nula, y no necesariamente que la diferencia en cuestión sea muy importante.

Por otra parte, el que no se rechace una H_0, no significa que se haya demostrado que dicha hipótesis nula es cierta, sino sólo que la misma es compatible con los datos observados, como lo serían probablemente también muchas otras hipótesis alternativas.

Autoevaluación: ¿qué interpretación práctica tiene la probabilidad 1-α asociada a un determinado intervalo de confianza?

El aserto probabilístico asociado a un determinado intervalo de confianza, como por ejemplo $P\left(\bar{x}-t_{N-1}(\alpha)\frac{s}{\sqrt{N}} < m < \bar{x}+t_{N-1}(\alpha)\frac{s}{\sqrt{N}}\right) = 1-\alpha$ debe interpretarse como que existe una probabilidad 1-α de que los dos límites del intervalo (que son aleatorios al depender de los parámetros muestrales $\bar{x}$ y s) comprendan entre ellos al verdadero valor (desconocido) de la media poblacional m, y no como que existe una probabilidad 1-α de que la media esté entre los dos valores obtenidos en un intervalo concreto

La media m no es una variable aleatoria, porque no existe una población a la que vayan asociados los posibles valores de m, y en la que una proporción 1-α de los individuos tengan valores de m comprendidos en dicho intervalo. Lo que sí que existe es una población de posibles muestras, a cada una de las cuales le corresponde un determinado intervalo de confianza, y en la que una proporción 1-α de las muestras llevan asociados intervalos de confianza que cubren al verdadero valor de m

Por tanto, cuando se dice (hablando con poca precisión) que existe una "probabilidad" 1-α de que un determinado intervalo contenga al verdadero valor de m, el término "probabilidad" no debe entenderse en el sentido técnico frecuencialista definido en este libro, sino en un sentido más coloquial como medida del grado de confianza que se puede tener en que la afirmación anterior es cierta.

Evaluación: se desea hacer una encuesta para estimar el % de votantes a los diferentes partidos en unas próximas elecciones. ¿Qué tamaño mínimo N debe tener la muestra si se desea que la probabilidad de cometer un error superior a 5 puntos porcentuales en la estimación de una determinada proporción P sea inferior al 1%?

Como $\frac{p^* - p}{\sqrt{p(1-p)/N}} = \sqrt{N}\frac{p^* - p}{\sqrt{p(1-p)}}$ se distribuye aproximadamente como una N(0,1) $\Rightarrow$

P(|p*-p|>0.05) = $P\left(\left|\sqrt{N}\frac{p^* - p}{\sqrt{p(1-p)}}\right| > 0.05\sqrt{N}/\sqrt{p(1-p)}\right) = P\left(|N(0,1)| > 0.05\sqrt{N}/\sqrt{p(1-p)}\right)$ ha de ser < 0.01

Mirando en la tabla de la normal se ve que para que se cumpla esto $0.05\sqrt{\frac{N}{p(1-p)}}$ ha de ser mayor que 2.575 y, por tanto N debe ser mayor que $\frac{2.575^2 \mathrm{p}(1-\mathrm{p})}{0.05^2} = 2652.25\mathrm{p}(1-\mathrm{p})$

Dado que p es desconocido, y teniendo en cuenta que p(1-p) nunca puede ser mayor que 0.25 (valor que se alcanzaría si p fuera igual a 0.5), en la expresión snterior se sustituye por ese valor, obteniéndose finalmente

N > 2625.25x0.25 = 663.1 ➔ N = 664 personas

8.A.2 Ejercicios adicionales

Suponiendo que los alumnos encuestados constituyen una muestra representativa de los jóvenes españoles, obtener un intervalo de confianza para la media y para la desviación típica de la estatura de los jóvenes varones en nuestro país.
Un estudio reciente afirma que los jóvenes españoles tienen una estatura media de 175 cms. con una desviación típica de 7 cms. ¿Son admisibles estas afirmaciones a la luz de la información contenida en la encuesta?

Hallar un intervalo de confianza para la altura media de las nubes que granizan (Datos en el archivo datgrani.sf3[1])
La variable analizada tiene una distribución asimétrica positiva, como puede comprobarse a partir de un gráfico en papel probabilístico. Comprobar que la asimetría es menos marcada si se realiza una transformación logarítmica de la variable y repetir el análisis trabajando con la variable transformada, hallando un intervalo de confianza para la mediana de la altura de las nubes que granizan.
Comparar las conclusiones obtenidas en los dos casos.

Para contrastar la hipótesis nula de que la varianza de una población normal es igual a 10, se toma de la misma una muestra de 25 datos, decidiéndose aceptar la H_0 si la varianza muestral resulta comprendida entre 5.75 y 15.17
a) Calcular el riesgo de 1ª especie α del procedimiento utilizado
b) ¿Qué riesgo β de 2ª especie tendrá dicho procedimiento si la varianza de la población es realmente igual a 15?

[1] Todos los archivos de datos que se mencionan en este texto pueden bajarse libremente de la URL http://personales.upv.es/rromero/descargas. Los alumnos de la Universidad Politécnica de Valencia pueden también bajárselos del Poliformat de la asignatura.

En una encuesta realizada sobre 80 empresas químicas españolas, 20 tenían certificado su sistema de calidad según la ISO 9001. Obtener un intervalo de confianza (con un nivel de confianza del 95%) para la proporción P de empresas químicas españolas que tienen certificado su sistema de calidad