Tema 5

Contraste de hipótesis estadísticas.

Contraste de hipótesis estadísticas

- Planteamiento del problema: hipótesis nula y alternativa.
- Tipos de errores.
- P-valor.
- Modelos de contrastes.
- Contrastes para una muestra.
- Contrastes para dos muestras.

Idea del contraste

- En ocasiones, nuestro objetivo **no es estimar un parámetro,** sino decidir decidir si puede considerarse “cierta” o “falsa” una proposición (hipótesis) acerca de la población.

- Los métodos que nos permitirán resolver este problema se denominan de **contrastes o tests de hipótesis**.

Contrastes parámetricos

- Los contrastes de hipótesis pueden ser de dos tipos:
  - Si la hipótesis formulada se refiere al valor que toma un parámetro de la distribución, se habla de **contrastes paramétricos**.
  - En caso contrario nos referimos a **contrastes no paramétricos.** En ellos se desconoce si las condiciones de la distribución de X o las de obtención de la muestra son las adecuadas para formular los modelos que resuelvan los problemas paramétricos anteriores.
- Por sencillez, nos centraremos en los **contrastes paramétricos.**

Toma de decisión con el contraste

- Los métodos de contraste (paramétricos) nos permiten, **a partir de la información obtenida de una muestra**, tomar con ciertas garantías **una decisión** sobre el valor de alguno de los parámetros de los que depende la distribución de la población utilizada.
- Trabajaremos con una(s) población(es) de distribución normal (o con tamaño de muestra $n \geq 30$) con parámetros desconocidos $\boldsymbol{\mu}$ y $\boldsymbol{\sigma}$, o con población(es) Bernoulli con **p** desconocido.

Formulación del contraste

- En los contrastes paramétricos:
  - Sobre el parámetro desconocido se plantea una hipótesis, **HIPÓTESIS NULA (H_o).**
  - El resto de posibilidades se engloban en la **HIPÓTESIS ALTERNATIVA (H_1).**

Regla de planteamiento

- Siempre que se plantee un contraste de hipótesis hay que tener en cuenta la siguiente regla:

  REGLA: **“La igualdad debe estar siempre en la hipótesis nula, nunca en la alternativa”.**

Ejemplo 1

- El 10% de las tarjetas de circuitos que produjo cierto fabricante durante un periodo reciente resultaron defectuosas. Se recomienda un cambio en el proceso de producción para tener una menor proporción de artículos defectuosos.
  - X = ¿La tarjeta de circuito producidas con el proceso modificado es defectuosa? ~ B(p), con p (desconocido) proporción real (poblacional) de tarjetas defectuosas que resultan del proceso modificado.
  - La hipótesis nula se establece en el hecho de que **el nuevo proceso no mejora la situación actual,** $H_0 : p \geq 0.10$ (→NO MODIFICO),
  - frente a que **el nuevo proceso mejora la proporción de defectuosos** $H_1 : p < 0.10$ (→ MODIFICO).

Ejemplo 2

- Queremos decidir sobre la compra de un coche. Para ello, me fijo en el ahorro medio a fin de mes, decidiendo la compra cuando es al menos de 500 €.
  - X = Ahorro mensual, en euros, observado al final de cada mes ~ $N(\mu,\sigma)$, con μ es el ahorro medio mensual (desconocido) y σ es la desviación típica (desconocida).
  - La hipótesis nula se establece en el hecho de que **el ahorro medio nos permite la compra del coche**, $H_0 : \mu \geq 500$ (→COMPRO),
  - frente a que **el ahorro medio NO nos permite la compra del coche**, $H_1 : \mu < 500$ (→NO COMPRO).

Ejemplo 3

- Cierto tipo de automóvil permanece sin daño visible el 25% de las veces en pruebas de colisión a 20 km/h. En un esfuerzo por incrementar este porcentaje se propone un diseño de defensa modificada.
  - X=¿El automóvil colisiona a 20 km/h sin daños visibles? ~ B(p), con p (desconocido) proporción real (poblacional) de colisiones a 20 km/h. sin daños visibles.
  - La hipótesis nula establece que **el nuevo proceso no mejora la situación actual**, $H_0 : p \leq 0.25$ (→NO MODIFICO),
  - frente a que **si mejora la proporción de vehículos sin daños visibles**, $H_1 : p > 0.25$ (→MODIFICO).

Ejemplo 4

- El tiempo de secado de cierto tipo de pintura tiene una distribución normal con valor medio de 75 minutos y desviación típica de 9 minutos. Se propone añadir un nuevo aditivo diseñado para disminuir el tiempo promedio de secado.
  - X = Tiempo, en minutos, de secado de cierto tipo de pintura con nuevo aditivo ~ N(μ,σ), donde μ es el tiempo medio de secado de la pintura con el aditivo (desconocido) y σ es la desviación típica (desconocida).
  - La hipótesis nula establece que **ese tiempo medio no aporta mejoras a la situación actual**, $H_0 : \mu \geq 75$ (→NO AÑADO),
  - frente a que, **es menor que el actual**, $H_1 : \mu < 75$ (→ AÑADO).

Ejemplo 5

- Un fabricante desea comparar el proceso de montaje habitual para uno de sus productos con un método propuesto que, supuestamente, reduce el tiempo de montaje. Se seleccionan 8 trabajadores de la planta de montaje y se les pide que monten las unidades con ambos procesos.
  - X_1 = Tiempo, en minutos, que tarda en montarse uno de los productos por el método habitual $\sim N(\mu_1,\sigma_1)$, donde μ_1 es el tiempo medio que un trabajador utiliza en montar un producto por el método habitual (desconocido) y σ_1 es la desviación típica (desconocida).
  - X_2 = Tiempo, en minutos, que tarda en montarse uno de los productos por el método propuesto $\sim N(\mu_2,\sigma_2)$, donde μ_2 es el tiempo medio que un trabajador utiliza en montar un producto por el método propuesto (desconocido) y σ_2 es la desviación típica (desconocida).

Ejemplo 5

- Para tomar la decisión de si, con el método propuesto, el tiempo medio de montaje del producto mejora contrastamos:
  - La hipótesis nula establece que **ese tiempo medio de montaje no aporta mejoras al obtenido con el método habitual**, $H_0 : \mu_1 \leq \mu_2$ (→NO CAMBIO EL MÉTODO),
  - frente a que, **el tiempo medio de montaje, con el nuevo método, es menor que el actual**, $H_1 : \mu_1 > \mu_2$ (→CAMBIO EL MÉTODO DE MONTAJE).

Ejemplo 6

- Dos empresas compiten en el sector de la producción de tornillos. Un cliente que lleva años trabajando con ambas empresas opina que el porcentaje de tornillos defectuosos producido por la fábrica A es igual que el de la fábrica B. Para comprobarlo se escoge una muestra de 100 tornillos de cada una de las fábricas, encontrando 9 defectuosos en A y 8 defectuosos en B.
  - X_1 = ¿Un tornillo producido por la fábrica A es defectuoso? $\sim B(p_1)$, con p_1 (desconocido) probabilidad de que un tornillo fabricado por A sea defectuoso (proporción real de tornillos defectuosos de la fábrica A).

Ejemplo 6

- X_2 = ¿Un tornillo producido por la fábrica B es defectuoso? ~ $B(p_2)$, con p_2 (desconocido) probabilidad de que un tornillo fabricado por B sea defectuoso (proporción real de tornillos defectuosos de la fábrica B).

- Para ver si el porcentaje de tornillos defectuosos producidos por ambas fábricas es el mismo o no contrastamos:
  - La hipótesis nula establece que **ese porcentaje es el mismo en ambos casos**, $H_0 : p_1 = p_2$ (→HAY IGUALDAD DE PROPORCIONES),
  - frente a que, **esos porcentajes son diferentes**, $H_1 : p_1 \neq p_2$ (→NO HAY IGUALDAD DE PROPORCIONES).

Información muestral

- Para tomar la decisión se tiene información aportada por **una muestra aleatoria simple (m.a.s) extraída de la(s) población(es)** X($_1$, X_2).
- A partir de los resultados de la(s) misma(s) se decide **rechazar la hipótesis nula** (aceptar la alternativa) o se decide **no rechazar la hipótesis nula**.

Errores de tipo I y II

- Tomando la decisión en base la información muestral (parte la población, pero no toda) puede ocurrir:

		Tomo la decisión de que la hipótesis correcta es…	
		Hipótesis Nula (H_0)	Hipótesis Alternativa (H_1)
En realidad, la hipótesis correcta es…	Hipótesis Nula (H_0)	CORRECTO	**ERROR TIPO I**
	Hipótesis Alternativa (H_1)	**ERROR TIPO II**	CORRECTO

Importancia del error de tipo I

- Parece razonable que para tomar la decisión busquemos que los errores sean lo más pequeños posibles (pero el valor del error no se puede computar ya que el valor real del parámetro es desconocido), hacemos pequeñas la probabilidad de que se cometa cada uno de dichos errores.
- **GARANTÍA:** El método que utilizamos para resolver el problema **limita la probabilidad de error tipo I a un valor pequeño** (entre 5 y 10%) y hace lo más pequeño posible la probabilidad de error tipo II. **El error tipo I se limita porque se considera el más grave, si ocurre.**
- Los ejemplos 7 y 8 pueden ilustrar esta afirmación.

Ejemplo 7

- Estamos ante la disyuntiva de APROBAR o SUSPENDER a un alumno en función de si sabe o no estadística. Se dispone de la nota de un examen (la muestra) para tomar la decisión:
  - Si la nota es 7 la decisión es aprobar.
  - Si la nota es 2 la decisión es suspender.
  - Si la nota es 4.6, ¿cuál es la decisión?

Ejemplo 7

- El contraste es H_0 : el alumno sabe (→LE APRUEBO)

  H_1 : el alumno no sabe (→LE SUSPENDO)
  - El **ERROR TIPO I** es suspender al alumno cuando tiene los conocimientos suficientes de la asignatura.
  - El **ERROR TIPO II** es aprobar al alumno cuando no tiene conocimientos suficientes de la asignatura.
- "Realmente" no sabemos los conocimientos del alumno y tomamos la decisión en base a un dato muestral (la nota del examen). En este caso, podemos cometer un ERROR TIPO I o un ERROR TIPO II con una determinada probabilidad.
- **Parece más grave cometer un ERROR TIPO I.**

Ejemplo 8

- En un juicio hay que decidir entre declarar inocente o culpable a una persona. El contraste es

  H_0 : La persona es inocente (→LA ABSUELVO)

  H_1 : La persona es culpable (→LA CONDENO)

- En este ejemplo, la muestra que se utiliza para tomar la decisión son las pruebas presentadas en el juicio. Los errores que pueden cometerse son:
  - ERROR TIPO I.- Condenar a un inocente.
  - ERROR TIPO II.- Absolver a un culpable.
- **Parece más grave cometer un ERROR TIPO I.**

Ejemplo 1

- El 10% de las tarjetas de circuitos que produjo cierto fabricante resultaron defectuosas. Se recomienda un cambio en el proceso de producción para tener una menor proporción de artículos defectuosos.
  - La hipótesis nula se establece en el hecho de que **el nuevo proceso no mejora la situación actual,** $H_0 : p \geq 0.10$ (→NO MODIFICO),
  - frente a que **el nuevo proceso mejora la proporción de defectuosos** $H_1 : p < 0.10$ (→ MODIFICO).
- ERROR TIPO I.- Modificamos el proceso de producción cuando no aporta ventajas frente al existente.
- ERROR TIPO II.- Mantenemos el proceso actual cuando podríamos mejorar la fabricación de tarjetas con el nuevo.

Ejemplo 2

- Compro coche si el ahorro medio a fin de mes es ≥ 500 €.
  - La hipótesis nula establece que **el ahorro medio nos permite la compra del coche**, $H_0 : \mu \geq 500$ (→COMPRO),
  - frente a que **el ahorro medio NO nos permite la compra del coche**, $H_1 : \mu < 500$ (→NO COMPRO).
- ERROR TIPO I.- Decidimos no comprar el coche cuando ahorramos lo suficiente para hacerlo.
- ERROR TIPO II.- Decidimos comprar el coche cuando no ahorramos lo suficiente para ello.
- En este caso, podemos plantearnos **la modificación de las hipótesis** ya que parece más grave el Error tipo II que el Error tipo I: $H_0 : \mu \leq 500$ y $H_1 : \mu > 500$.

Ejemplo 3

- En un automóvil que permanece sin daño visible el 25% de las veces se propone un diseño de defensa modificada para incrementar este porcentaje
  - La hipótesis nula establece que **el nuevo proceso no mejora la situación actual**, $H_0 : p \leq 0.25$ (→NO MODIFICO),
  - frente a que **si mejora la proporción de vehículos sin daños visibles**, $H_1 : p > 0.25$ (→MODIFICO).
- ERROR TIPO I.- Decidimos MODIFICAR el proceso cuando el nuevo proceso no mejora.
- ERROR TIPO II.- Decidimos NO MODIFICAR el proceso cuando el nuevo proceso mejora.

Ejemplo 4

- El tiempo de secado de cierto tipo de pintura tiene una distribución normal con valor medio de 75 minutos y desviación típica de 9 minutos. Se propone <u>añadir un nuevo aditivo diseñado para disminuir el tiempo promedio de secado</u>.
  - La hipótesis nula establece que **ese tiempo medio no aporta mejoras a la situación actual**, $H_0 : \mu \geq 75$ ($\rightarrow$NO AÑADO),
  - frente a que, **es menor que el actual**, $H_1 : \mu < 75$ ($\rightarrow$ AÑADO).
- ERROR TIPO I.- Decidimos AÑADIR el aditivo cuando el tiempo medio de secado no se acorta.
- ERROR TIPO II.- Decidimos NO AÑADIR el aditivo cuando el tiempo medio de secado se acorta.

Ejemplo 5

- Un fabricante desea comparar el proceso de montaje habitual para uno de sus productos con un método propuesto que, supuestamente, reduce el tiempo de montaje.
  - La hipótesis nula establece que **ese tiempo medio de montaje no aporta mejoras al obtenido con el método habitual**, $H_0 : \mu_1 \leq \mu_2$ (→NO CAMBIO EL MÉTODO),
  - frente a que, **el tiempo medio de montaje, con el nuevo método, es menor que el actual**, $H_1 : \mu_1 > \mu_2$ (→CAMBIO EL MÉTODO DE MONTAJE).
- ERROR TIPO I.- Decidimos CAMBIAR de método de trabajo cuando el nuevo no aporta mejoras al método habitual.
- ERROR TIPO II.- Decidimos NO CAMBIAR de método de trabajo cuando el nuevo método aporta mejoras al método habitual.

Ejemplo 6

- Un cliente opina que el porcentaje de tornillos defectuosos producido por la fábrica A es igual que el de la fábrica B.
  - La hipótesis nula establece que **ese porcentaje es el mismo en ambos casos**, $H_0 : p_1 = p_2$ (→HAY IGUALDAD DE PROPORCIONES),
  - frente a que, **esos porcentajes son diferentes**, $H_1 : p_1 \neq p_2$ (→NO HAY IGUALDAD DE PROPORCIONES).
- ERROR TIPO I.- Decidimos que HAY DIFERENCIAS SIGNIFICATIVAS entre ambos porcentajes cuando en realidad no son distintos los porcentajes de ambos productores
- ERROR TIPO II.- Decidimos NO HAY DIFERENCIAS SIGNIFICATIVAS entre ambos porcentajes cuando en realidad sí son distintos los porcentajes de ambos productores.

Nivel de significación del contraste

- Para unos valores de la muestra rechazaremos la hipótesis nula mientras que para otros valores no rechazaremos la hipótesis nula.
- El conjunto de valores para los que rechazamos la hipótesis nula se conoce como **región crítica del contraste,** el complementario es **la región de aceptación** de la hipótesis nula.
- Se llama **nivel de significación del contraste** a

α = P(ERROR TIPO I)=P(Rechazar H_0/H_0 es cierta) = P(Región crítica /H_0 es cierta).

P-valor

- Para tomar la decisión de rechazar o no rechazar la hipótesis nula planteada utilizaremos el llamado **nivel crítico (o p-valor) del test o contraste**, que es **el menor nivel de significación** a partir del cual la hipótesis nula comienza a ser rechazada.
- **Interpretación:** el p-valor se puede entender como el grado de compatibilidad de los datos con la hipótesis nula.

Procedimiento del contraste (I)

- ¿Cuándo rechazamos la hipótesis nula planteada sobre el parámetro? Cuando el **resultado muestral es muy diferente** de la hipótesis que realizamos sobre el parámetro.
- **Pensemos en el Ejemplo 4:**
  - X = Tiempo, en minutos, de secado de cierto tipo de pintura con nuevo aditivo ~ $N(\mu,\sigma)$, donde μ es el tiempo medio de secado de la pintura con el aditivo (desconocido) y σ es la desviación típica (desconocida).
  - **Tenemos una m.a.s. de 25 tiempos de secado. (X_1, X_2,…, X_{25}) de la población, con**

$$\overline{X} = 70.8 \text{ min.} \quad \text{y} \quad S = 9 \text{ min.}$$

Procedimiento del contraste (II)

- El contraste es: $H_0 : \mu \geq 75$ ($\rightarrow$NO AÑADO ADITIVO)
  $H_1 : \mu < 75$ ($\rightarrow$ AÑADO ADITIVO)
- Si queremos contrastar si $\mu \geq 75$ comprobaré si el correspondiente valor muestral (media muestral) es mayor o igual que 75 y en ese caso no rechazaré H_0. Pero se intuye que debemos no rechazar la hipótesis nula como <u>cierta incluso si estuviera un poco por debajo de 75</u> (ya que, si la hipótesis nula fuese cierta, sería posible que la media de una muestra de 25 datos estuviese un poco por debajo de ese valor debido al azar). Por supuesto, <u>cuanto más pequeña sea la media</u> muestral y más alejada esté de 75, menos incertidumbre tendremos para tomar la decisión de que la hipótesis nula es falsa.

Procedimiento del contraste (III)

- ¿Hasta qué valor de la media muestral por debajo de 75, es razonable seguir sosteniendo que la hipótesis nula es cierta? O ¿a partir de qué valor de la media muestral es razonable pensar que la hipótesis nula NO es cierta? **En concreto, para el valor obtenido en nuestra muestra**

$$\overline{X} = 70.8 \text{ min}.$$

¿es más razonable no rechazar la hipótesis nula o lo es rechazarla?

Procedimiento del contraste (IV)

- Sabemos que :

$$T = \frac{\overline{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}.$$

- En nuestro ejemplo, si la hipótesis nula es cierta, X~ N(75;σ). Por tanto, para la muestra de 25 tiempos de secado y cuando la hipótesis nula es cierta el valor del estadístico es

$$T = \frac{70.8 - 75}{9/\sqrt{25}} = -2.3333.$$

- Rechazaremos H_o: $\mu \geq 75$ cuanto menor sea el valor del estimador media muestral (y por tanto de T), aunque queda por determinar el cuánto.

Procedimiento del contraste (V)

- Consideramos que vamos a rechazar la hipótesis nula cuando el valor del estadístico T es -2.3333 o más pequeño:

$$P\left(\overline{X} \leq 70.8 \big/ H_0 \text{ cierta}\right) = P\left(T = \frac{\overline{X} - \mu}{S \big/ \sqrt{n}} \leq -2.3333\right) =$$

$$= P(t_{24} \leq -2.3333) = 0.0142.$$

- Esta probabilidad recibe el nombre de **Nivel crítico (o p-valor) del test o contraste**, y es el menor nivel de significación a partir del cual la hipótesis nula comienza a ser rechazada.

Procedimiento del contraste (VI)

Esta probabilidad nos mide si es creíble, cuando es cierta Ho, que el valor del estadístico alcance el valor -2.3333 o menor: **se puede interpretar que si rechazásemos la hipótesis nula con el valor del estadístico T menor o igual que -2.3333 cometeríamos el error tipo I en un 1.42% de las ocasiones (es decir, nos equivocaríamos pocas veces). Por tanto, parece lógico que en este caso rechacemos la hipótesis nula porque, con esta decisión, nos equivocamos en pocas ocasiones. <u>Intuitivamente, rechazaremos la hipótesis nula cuando esta probabilidad sea muy pequeña (cuando nos equivocamos, cometiendo el error tipo I, un porcentaje pequeño de ocasiones).</u>**

Valores del P-valor

- Si trabajamos como mucho con un 5% de error tipo I, **no rechazamos la hipótesis nula** si el p-valor es mayor o igual que 0.05, **rechazando la hipótesis nula** si el p-valor es menor de 0.05.
- Si trabajamos como mucho con un 10% de error tipo I, **no rechazamos la hipótesis nula** si el p-valor es mayor o igual que 0.10, **rechazando la hipótesis nula** si el p-valor es menor de 0.10.
- Para valores entre el 5 y 10% (entre 0.05 y 0.10 de probabilidad de error tipo I o nivel de significación) se actúa de forma similar.
- **No suele trabajarse con probabilidades de error tipo I, o nivel de significación, mayores.**

Ejemplo 4

- En el ejemplo, el p-valor es 0.0142. ¿Qué hacemos? Con los datos del ejemplo 4, como $0.0142 < 0.05 < 0.1$, se puede **tomar la decisión de rechazar la hipótesis** para cualquiera de los niveles de significación habituales → **AÑADO EL ADITIVO**.
- **<u>Modificamos datos del problema 4:</u>** ¿Qué decisión tomaríamos si en la muestra de 25 datos de secado hubiéramos obtenido **un tiempo medio de secado de 75 minutos** con una desviación de 9 minutos?

Ejemplo 4

- Con esta modificación los cálculos quedarían:

$$P\left(\overline{X} \leq 75 \Big/ H_0 \text{ cierta}\right) = P\left(T = \frac{\overline{X} - \mu}{S \big/ \sqrt{n}} \leq \frac{75 - 75}{9 \big/ \sqrt{25}} = 0\right) =$$

$$= P(t_{24} \leq 0) = 0.5.$$

- **Es decir, si rechazásemos la hipótesis nula con T ≤ 0 cometeríamos el error tipo I en un 50% de las ocasiones (es decir, nos equivocaríamos muchas veces). Por tanto, NO rechazaremos Ho → NO AÑADO EL ADITIVO**.
- Con estos datos en el ejemplo, como <u>0.5 > 0.1 >0.05, se toma la decisión de **no rechazar la hipótesis nula** para cualquier nivel de significación habitual</u>.

Modelos de contrastes

- Pueden presentarse 3 modelos distintos de contraste de hipótesis:

$$\begin{cases} H_0 : \theta = \theta_0 \\ H_1 : \theta \neq \theta_0 \end{cases} (\text{mod.}1) \quad \begin{cases} H_0 : \theta \geq \theta_0 \\ H_1 : \theta < \theta_0 \end{cases} (\text{mod.}2) \quad \begin{cases} H_0 : \theta \leq \theta_0 \\ H_1 : \theta > \theta_0 \end{cases} (\text{mod.}3)$$

- La resolución de los 3 modelos de contrastes puede plantearse de forma sistemática singularizando la diferencia que hay entre ellos.

Modelo 1 $\begin{cases} H_0 : \theta = \theta_0 \\ H_1 : \theta \neq \theta_0 \end{cases}$ (mod. 1)

- El método de decisión tiene los siguientes pasos:

1. Se elige el estadístico T con el que trabajar y del que conocemos como se comporta (conocemos su distribución de probabilidad): **T~ distribución conocida**.
2. Se supone que la hipótesis nula es cierta, entonces θ se sustituye por θ_0 y además también se sustituyen los correspondientes valores muestrales: **T= t_o = número real.**
3. Paso 3.- Se calcula el nivel crítico:

$$\alpha_0 = 2 \cdot P\left(T \geq \left|t_0\right|\right).$$

4. Se rechaza la H_0 si $\alpha_0 < 0.05$, no se rechaza en caso contrario.

Modelo 2 $\begin{cases} H_0 : \theta \geq \theta_0 \\ H_1 : \theta < \theta_0 \end{cases}$ (mod. 1)

- El método de decisión tiene los siguientes pasos:

1. Se elige el estadístico T con el que trabajar y del que conocemos como se comporta (conocemos su distribución de probabilidad): **T~ distribución conocida**.
2. Se supone que la hipótesis nula es cierta, entonces θ se sustituye por θ_0 y además también se sustituyen los correspondientes valores muestrales: **T= $\mathbf{t_o}$ = número real.**
3. Paso 3.- Se calcula el nivel crítico:

$$\alpha_0 = P(T \leq t_0).$$

4. Se rechaza la H_0 si α_0<0.05, no se rechaza si α_0<0.05, no se rechaza en caso contrario.

Modelo 3 $\begin{cases} H_0 : \theta \leq \theta_0 \\ H_1 : \theta > \theta_0 \end{cases}$ (mod. 1)

- El método de decisión tiene los siguientes pasos:

1. Se elige el estadístico T con el que trabajar y del que conocemos como se comporta (conocemos su distribución de probabilidad): **T~ distribución conocida**.
2. Se supone que la hipótesis nula es cierta, entonces θ se sustituye por θ_0 y además también se sustituyen los correspondientes valores muestrales: **T= t_o = número real.**
3. Paso 3.- Se calcula el nivel crítico:

$$\alpha_0 = P(T \geq t_0).$$

4. Se rechaza la H_0 si $\alpha_0 < 0.05$, no se rechaza en caso contrario.

Tipos de contrastes con una muestra

1. **Contraste de hipótesis para la <u>media</u> de una población normal X~N(μ, σ) (o de cualquier población con tamaño de muestra n≥30) con desviación típica poblacional desconocida.**

Se utiliza el estadístico $T = \dfrac{\overline{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}.$

2. **Contraste de hipótesis para la <u>proporción de éxitos</u> p de una población Bernoulli, X~B(p) .**

Se utiliza el estadístico $T = \dfrac{\hat{p} - p}{\sqrt{\dfrac{p(1-p)}{n}}} \sim N(0,1).$

Problema 1

El tiempo de secado de cierto tipo de pintura en condiciones de prueba específicas, tiene una distribución normal con valor medio de 75 minutos. Se propone añadir un nuevo aditivo diseñado para disminuir el tiempo medio de secado. Se cree que los tiempos de secado con este aditivo aún tienen una distribución normal. Se toma una muestra de 20 tiempos de secado obteniendo una media de 72 minutos con una desviación de 8.35 minutos

a) El promotor del cambio de mezcla, sostiene que con el nuevo producto disminuye el tiempo de secado medio.
b) Otro Ingeniero de la empresa sostiene que el compañero que propuso la mezcla puede tener razón pero que tiene dudas para afirmar, tajantemente, que mejora pero cree que, al menos, no empeora el tiempo medio actual.
c) Un tercer Ingeniero sostiene que con el nuevo aditivo nada cambia y se mantiene el mismo tiempo medio de secado.

Problema 2

El 10% de las tarjetas de circuitos que produjo cierto fabricante durante un periodo reciente resultaron defectuosas. Un ingeniero recomienda un cambio en el proceso de producción, con la creencia de que dará como resultado una menor proporción de artículos defectuosos. En una muestra de 250 tarjetas producidas con el proceso renovado se tienen 20 tarjetas defectuosas

a) Otro ingeniero defiende que el nuevo proceso no aporta ninguna novedad respecto al existente hasta este momento. Explica la decisión que tomará respecto a la falta de novedad en la recomendación.
b) Para comprobar si mejoró el proceso de fabricación, el Ingeniero que propuso la modificación utiliza los datos de la muestra. Explica la decisión que tomará respecto a la mejora del proceso de fabricación.
c) Un tercer compañero sostiene que tiene dudas sobre las ventajas del nuevo proceso aunque sostiene, que al menos, no empeora el porcentaje de tarjetas defectuosas actual. Explica la decisión que se tomará.

Tipos de contrastes con dos muestras I

1. **Contraste de hipótesis para la diferencia de medias de poblaciones normales (o con tamaños muestrales al menos de 30) independientes:**
   A. **Las desviaciones típicas son desconocidas y no se hacen suposiciones sobre ellas.**
   B. **Las desviaciones típicas son desconocidas pero iguales.**
2. **Contraste de hipótesis para la diferencia de medias de poblaciones normales (o con tamaños muestrales al menos de 30) de datos pareados.**

Tipos de contrastes con dos muestras II

3. **Contraste de hipótesis para la <u>diferencia de las proporciones</u> de éxitos de dos poblaciones Bernoulli.**
4. **Contraste de hipótesis para el <u>cociente de varianzas</u> de poblaciones normales (o con tamaños muestrales al menos de 30).**

Contrastes tipo 1. A

- **Si realizo un contraste de hipótesis para la <u>diferencia de medias</u> ($\theta=\mu_1-\mu_2$, $\theta_0=0$) de dos poblaciones normales <u>independientes</u> $X_1 \sim N(\mu_1, \sigma_1)$ y $X_2 \sim N(\mu_2, \sigma_2)$ (o con tamaños muestrales al menos de 30) en el caso de que las desviaciones típicas son desconocidas (y no se hacen suposiciones sobre ellas)** utilizo el estadístico:

$$T = \frac{\left(\overline{X}_1 - \overline{X}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}} \sim t_f, \quad \text{con } f = \frac{\left(\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}\right)^2}{\dfrac{\left(\dfrac{S_1^2}{n_1}\right)^2}{n_1 - 1} + \dfrac{\left(\dfrac{S_2^2}{n_2}\right)^2}{n_2 - 1}}.$$

Contrastes tipo 1. B

- **Si realizo un contraste de hipótesis para la <u>diferencia de medias</u> ($\theta=\mu_1-\mu_2$, $\theta_0=0$) de poblaciones normales <u>independientes</u> $X_1\sim N(\mu_1, \sigma)$ y $X_2\sim N(\mu_2, \sigma)$ (o con tamaños muestrales al menos de 30) en el caso de que las desviaciones típicas sean desconocidas pero iguales** utilizo el estadístico:

$$T=\frac{\left(\overline{X}_1-\overline{X}_2\right)-\left(\mu_1-\mu_2\right)}{\sqrt{\dfrac{S_p^2}{n_1}+\dfrac{S_p^2}{n_2}}}\sim t_{n_1+n_2-2}, \quad \text{con } S_p^2=\frac{\left(n_1-1\right)S_1^2+\left(n_2-1\right)S_2^2}{n_1+n_2-2}.$$

Contrastes tipo 2

- **Si X_1~N(μ_1, σ_1) y X_2~N(μ_2, σ_2) (o con tamaños muestrales al menos de 30)** siendo **$D=X_1-X_2$. En un contraste de hipótesis para la <u>diferencia de medias</u> ($\theta=\mu_1-\mu_2=\mu_D$, $\theta_0=0$) de poblaciones normales <u>correspondientes a datos pareados</u>** utilizo el estadístico:

$$T = \frac{\overline{D} - \mu_D}{S_D / \sqrt{n}} \sim t_{n-1}.$$

Contraste para la diferencia de p

- **Si realizo un contraste de hipótesis para la <u>diferencia de las proporciones</u> ($\theta=p_1-p_2$ y $\theta_0=0$) de éxitos de dos poblaciones Bernoulli, $X_1\sim B(p_1)$ y $X_2\sim B(p_2)$,** utilizo el estadístico:

$$T=\frac{(\hat{p}_1-\hat{p}_2)-(p_1-p_2)}{\sqrt{\frac{\hat{p}_1\cdot(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2\cdot(1-\hat{p}_2)}{n_2}}}\sim N(0,1).$$