

HERRAMIENTAS ESTADÍSTICAS -COMPARACIÓN DE DOS MUESTRAS: t-test

Módulo 12

APUNTES DE CLASE

Profesor: Arturo Ruiz-Falcó Rojas

Madrid, Mayo 2009

ÍNDICE

4.	¿QUÉ ES UN CONTRASTE DE HIPÓTESIS?	3
5.	METODOLOGÍA DEL CONTRASTE DE HIPÓTESIS.....	3
6.	CONTRASTES EN POBLACIONES NORMALES DE USO FRECUENTE.....	7
7.	COMPARACIÓN DE DOS TRATAMIENTOS	8
7.1.	EL PROCESO DE MEJORA MÁS SIMPLE: COMPARACIÓN DE ALTERNATIVAS.....	8
7.2.	COMPARACIÓN DE LA MEDIA DE DOS PROCESOS DE FABRICACIÓN. CASO VARIANZA CONOCIDA	9
7.3.	COMPARACIÓN DE LA MEDIA DE DOS PROCESOS DE FABRICACIÓN. CASO VARIANZA DESCONOCIDA	13
7.4.	FORMACIÓN DE BLOQUES ALEATORIZADOS: CASO DE ENSAYOS EMPAREJADOS. VENTAJAS E INCONVENIENTES	15
8.	CONTRASTES DE AJUSTE	19
8.1.	Concepto de contraste de ajuste	19
8.2.	Contraste Chi-cuadrado de Pearson.....	20
8.3.	Contraste de Kolmogorov Smirnov (KS)	23
8.4.	Contraste de Anderson-Darling (AD)	24
8.5.	Aplicación a los contrastes de normalidad	24
9.	BATERÍA DE EJERCICIOS DE CONTRASTES	27
9.1.	CASO NOVOMAT	30
9.1.1.	Escena 1	30
9.1.2.	Escena 2.....	31
9.1.3.	Escena 3.....	31

4. ¿QUÉ ES UN CONTRASTE DE HIPÓTESIS?

Un contraste estadístico de hipótesis consiste en evaluar una población a través de una muestra aleatoria de la misma, con objeto de determinar si existen evidencias para que una hipótesis formulada sobre la población deba ser rechazada o no pueda ser rechazada.

Según su aplicación, se pueden dividir los contrastes:

- Contrastes de un valor concreto de un parámetro del modelo.
- Contrastes para validar la función de distribución que se supone ha generado los datos (contrastos de ajuste).

De momento nos referiremos únicamente a los contrastes del primer tipo, dejando los contrastes de ajuste para un epígrafe monográfico al final de este módulo.

5. METODOLOGÍA DEL CONTRASTE DE HIPÓTESIS

Un contraste de hipótesis requiere naturalmente el establecimiento de una hipótesis previa, que se denomina hipótesis nula y se representa por H_0 , y compararla con otra, que se denomina hipótesis alternativa y se representa por H_1 . Puede ocurrir que las hipótesis formuladas sean simples o compuestas. Hipótesis simples son aquellas que presuponen un valor dado para el parámetro y compuestas son aquellas que no son simples. Por ejemplo, si se formulan las hipótesis:

- H_0 : La probabilidad de sacar cara al lanzar una moneda es $p=0.5$.
- H_1 : La probabilidad de sacar cara al lanzar una moneda es $p \neq 0.5$.

H_0 , es una hipótesis simple y H_1 es una hipótesis compuesta.

Una vez fijada la hipótesis sobre un parámetro de la población τ , es preciso encontrar un estadístico pivote, que esté relacionado con este parámetro y cuya

HERRAMIENTAS ESTADÍSTICAS-COMPARACIÓN DE DOS MUESTRAS: t-test

distribución sea conocida. Este estadístico va a aportar el criterio de decisión según el valor que alcance en la muestra. Para fijar el valor crítico, frontera que marca el cambio de decisión, es necesario fijar el riesgo de equivocarse que se está dispuesto a asumir. En efecto, en un contraste de hipótesis se pueden cometer errores de dos tipos:

- Error tipo I: Consiste en rechazar la hipótesis nula cuando es cierta.
- Error tipo II: Consiste en aceptar la hipótesis nula cuando es falsa.

DECISIÓN TOMADA	VERDAD DESCONOCIDA	
	H_0	H_1
NO SE RECHAZA H_0 ("SE ACEPTA H_0 ")	ACIERTO $p = 1 - \alpha$	ERROR TIPO II $p = \beta$
SE RECHAZA H_0 ("SE ACEPTA H_1 ")	ERROR TIPO I $p = \alpha$	ACIERTO $p = 1 - \beta$

Tabla 1: Errores en un contraste de hipótesis

A la probabilidad α de rechazar la hipótesis nula siendo cierta, se le denomina nivel de significación del contraste, a la probabilidad $1 - \beta$ de rechazar la hipótesis nula siendo falsa se le denomina potencia del contraste. β es función del valor real de τ de la población, es decir, $\beta = \beta(\tau)$. A la función $\beta = \beta(\tau)$ se le llama característica de operación o simplemente curva CO y a la función $\Gamma(\tau) = 1 - \beta(\tau)$ se le llama función de potencia del contraste.

Como puede observarse existe un paralelismo con la elaboración de intervalos de confianza.

El modo de calcular el valor crítico c varía según estén formuladas H_0 y H_1 (ver Figura 1).

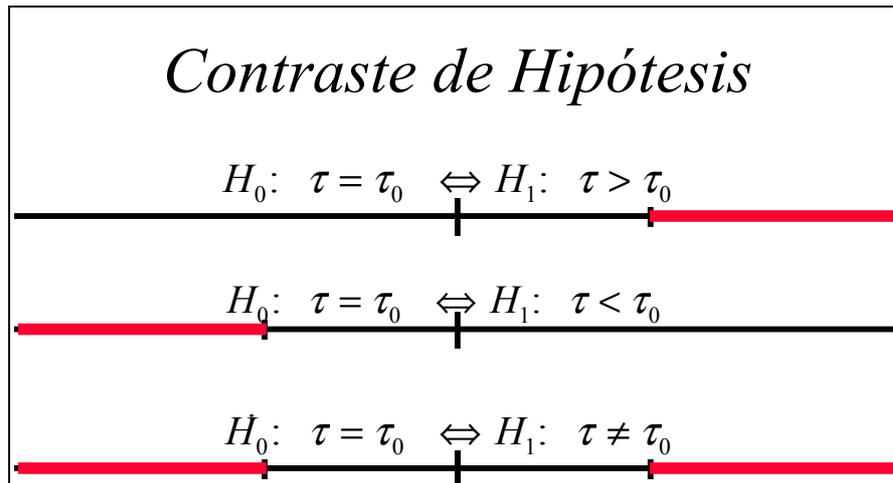


Figura 1: Valor crítico c

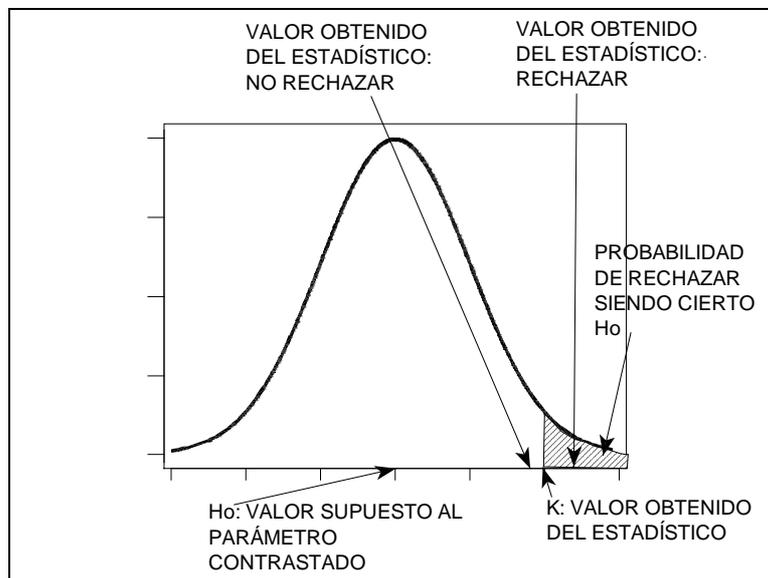


Figura 2: Distribución del estadístico pivote

EJEMPLO 12. ¿SE HA DESPLAZADO EL PROCESO?

Una máquina herramienta produce unas piezas cuya calidad se caracteriza, entre otros factores, por el diámetro exterior. Por datos históricos se sabe que la

HERRAMIENTAS ESTADÍSTICAS-COMPARACIÓN DE DOS MUESTRAS: t-test

desviación típica del proceso de mecanizado sensiblemente constante e igual a 0,30 mm. La máquina cuenta con un sistema de ajuste que permite variar el diámetro del mecanizado. El diámetro nominal es 9,75 mm siendo las tolerancias simétricas. Para comprobar si la máquina está correctamente ajustada, se mide una muestra aleatoria de cinco piezas del lote. Los resultados de la medida son 9,40 mm, 9,70 mm, 9,65 mm, 8,95 mm y 9,50 mm ¿Está el proceso de mecanizado ajustado en 9,75 mm?

En el capítulo anterior vimos que la media muestral se distribuye según una $N(\mu, \sigma/\sqrt{n})$. Entonces el estadístico:

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

se distribuirá según una $N(0,1)$. En este caso $\mu = 9,75$, $\sigma = 0,3$, $\bar{x} = 9,44$ y $n = 5$. Con estos datos resulta:

$$z = \frac{9,44 - 9,75}{0,3 / \sqrt{5}} = -2,3$$

Por las tablas de la distribución normal sabemos que, si efectivamente la máquina está trabajando centrada en 9,75 mm, tenemos aproximadamente un 2% de probabilidad de sacar una muestra de tamaño 5 cuyo diámetro medio se aleje del nominal en más de 0.31 mm. ¿Son entonces los resultados de la muestra una evidencia de que el proceso está desajustado? La estadística no es capaz de dar una respuesta contundente, únicamente dice que con un nivel de significación inferior al 1% se puede rechazar esta hipótesis. ¿Cuál es el nivel de significación que hay que tomar? A esa pregunta no la responde la estadística sino el sentido común de quien utiliza la estadística. Solo diremos que son habituales valores de α del 5% o del 10%. Por lo tanto si adoptamos un nivel de significación superior al 2% deberemos rechazar la hipótesis que la máquina está ajustada y proceder a su reajuste.

APLICACIÓN MINITAB

Para resolver este problema, Minitab dispone de la opción **Stat->Basic Statistics->1-Sample z**. Se obtiene:

One-Sample Z: Diámetros mm

HERRAMIENTAS ESTADÍSTICAS-COMPARACIÓN DE DOS MUESTRAS: t-test

Test of mu = 9,75 vs not = 9,75
The assumed standard deviation = 0,3

Variable	N	Mean	StDev	SE Mean	95% CI	Z	P
Diámetros mm	5	9,44000	0,29875	0,13416	(9,17704; 9,70296)	-2,31	0,021

6. CONTRASTES EN POBLACIONES NORMALES DE USO FRECUENTE.

Los siguientes contrastes son muy clásicos en la literatura estadística:

- Contraste de la media μ siendo σ conocido.
- Contraste de la media μ siendo σ desconocido.
- Contraste de la varianza σ^2 .
- Contraste de igualdad de varianzas $\sigma_1^2 = \sigma_2^2$ de dos poblaciones.
- Contraste de la diferencia de medias de dos poblaciones $\mu_1 - \mu_2$ siendo las varianzas iguales $\sigma_1^2 = \sigma_2^2$.

En la tabla siguiente se indican los estadísticos pivote y su distribución.

REF.	HIPÓTESIS	PARÁMETRO	ESTADÍSTICO	DISTRIBUCIÓN
1	$H_0 : \mu = \mu_0$ $H_1 : \mu \neq \mu_0$	μ (σ CONOCIDO)	$\frac{(\bar{x} - \mu_0)}{\sigma/\sqrt{n}}$	$N(0,1)$
2	$H_0 : \mu = \mu_0$ $H_1 : \mu \neq \mu_0$	μ (σ DESCONOCIDO)	$\frac{(\bar{x} - \mu_0)}{s/\sqrt{n}}$	t_{n-1}
3	$H_0 : \sigma^2 = \sigma_0^2$ $H_1 : \sigma^2 \neq \sigma_0^2$	σ^2	$\frac{ns^2}{\sigma_0^2}$	χ^2_{n-1}
4	$H_0 : \sigma_1^2 = \sigma_2^2$	σ^2	$\frac{s_1^2}{s_2^2}$	$F_{n-1,m-1}$

HERRAMIENTAS ESTADÍSTICAS-COMPARACIÓN DE DOS MUESTRAS: t-test

REF.	HIPÓTESIS	PARÁMETRO	ESTADÍSTICO	DISTRIBUCIÓN
	$H_1 : \sigma_1^2 \neq \sigma_2^2$			
5	$H_0 : \mu_1 = \mu_2$ $H_1 : \mu_1 \neq \mu_2$	μ (σ IGUALES) (σ CONOCIDO)	$\frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)}}$	$N(0, 1)$
6	$H_0 : \mu_1 = \mu_2$ $H_1 : \mu_1 \neq \mu_2$	μ (σ IGUALES) (σ DESCONOCIDO)	$\frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$	t_{m+n-2}

Tabla 2: Contrastes en poblaciones normales de uso frecuente

. Para que las conclusiones sean válidas, es preciso que:

- Las muestras sean aleatorias y por tanto representen fielmente la alternativa que representan.
- Que los datos de partida sean normales. En el caso de los contrastes de la media este aspecto no es crítico si la muestra es suficientemente grande (20-30 unidades, aunque depende del grado de no normalidad).

7. COMPARACIÓN DE DOS TRATAMIENTOS

7.1. EL PROCESO DE MEJORA MÁS SIMPLE: COMPARACIÓN DE ALTERNATIVAS

Una vez identificada una alternativa a la situación actual de un proceso, el proceso de mejora pasa por la comparación de ambas alternativas para ver si realmente la nueva alternativa es mejor que la existente. En esta comparación es preciso

distinguir aquellas diferencias que realmente significan que una alternativa es superior a la otra, de las fluctuaciones aleatorias.

7.2. COMPARACIÓN DE LA MEDIA DE DOS PROCESOS DE FABRICACIÓN. CASO VARIANZA CONOCIDA

EJEMPLO 13. COMPARACIÓN DE LA RESISTENCIA AL DESGASTE DE SUELAS DE ZAPATOS

Un fabricante de zapatos está estudiando un material alternativo para las suelas. La resistencia al desgaste es la característica de calidad más importante. Para ello se ha diseñado un experimento en el que a cinco niños (denominado "grupo A") se les ha calzado con zapatos con la suela actual y a otros cinco niños (denominado "grupo B") se les ha calzado con zapatos con la nueva suela. Mediante un procedimiento normalizado, se mide el espesor de la suela de cada uno de los 20 zapatos y al cabo de un mes se vuelve a medir de manera que es posible calcular el desgaste de cada zapato.

Este fabricante cuenta con un histórico de datos que le permite suponer que la desviación típica del desgaste es $\sigma = 0,35$ mm.

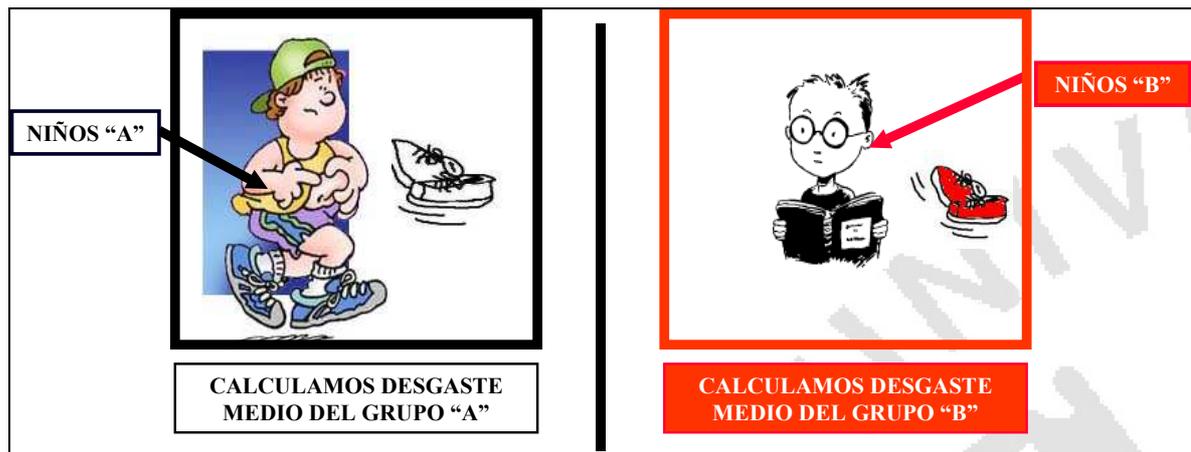


Figura 3: Ensayo no pareado

El planteamiento es:

HERRAMIENTAS ESTADÍSTICAS-COMPARACIÓN DE DOS MUESTRAS: t-test

- H_0 : Ambos materiales tienen la misma resistencia al desgaste: $\mu_A = \mu_B$.
- H_1 : Ambos materiales **no** tienen la misma resistencia al desgaste (uno de los dos es peor): $\mu_A \neq \mu_B$.

El estadístico pivote a emplear en este caso es el caso 5 de la Tabla 2. Antes de realizar el contraste y como rutina se debe hacer una representación gráfica de los datos y calcular los estadísticos de las muestras. Los resultados obtenidos fueron:

Grupo B	Grupo A
0,46	0,46
0,62	0,71
0,62	0,67
0,77	0,51
1,11	0,8
0,47	0,68
0,66	0,66
0,57	1,15
0,61	0,57
0,82	0,84

Tabla 3: Resultados obtenidos del desgaste de los zapatos

APLICACIÓN MINITAB

En Minitab esto se hace:

Graf->Dotplot

Stat->Basic Statistics->Display descriptive statistics

HERRAMIENTAS ESTADÍSTICAS-COMPARACIÓN DE DOS MUESTRAS: t-test

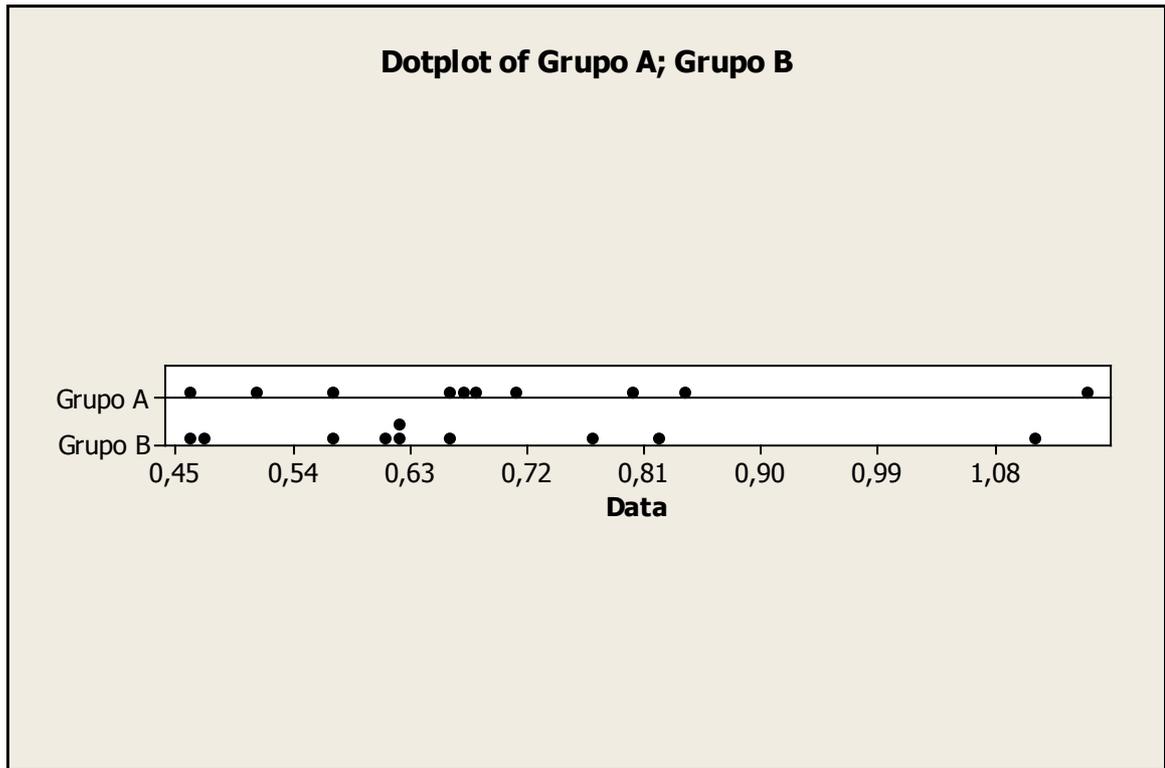


Figura 4: Diagrama de puntos del desgaste de los zapatos

Descriptive Statistics: Grupo A; Grupo B

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3
Grupo A	10	0	0,7033	0,0618	0,1954	0,4603	0,5509	0,6736	0,8086
Grupo B	10	0	0,6698	0,0609	0,1926	0,4592	0,5432	0,6163	0,7818

Variable	Maximum
Grupo A	1,1460
Grupo B	1,1123

Estudiando los resultados (ver Figura 4) se ve que:

- Los puntos se distribuyen de una manera similar, estando dos de ellos separados del resto.
- Las desviaciones típicas son bastante parecidas y son menores que la supuesta $\sigma = 0,35$ mm.

HERRAMIENTAS ESTADÍSTICAS-COMPARACIÓN DE DOS MUESTRAS: t-test

- El desgaste medio de los zapatos de los niños del grupo B es menor. Es preciso analizar si esta diferencia es significativa (caso 5 de la Tabla 2).

$$z = \frac{(\bar{x}_A - \bar{x}_B) - (\mu_A - \mu_B)}{\sqrt{\left(\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}\right)}} = \frac{0,70 - 0,67}{\sqrt{\left(\frac{2 * 0,35^2}{10}\right)}} = 0,19$$

El valor crítico del estadístico con $\alpha=0,05$ es $z_{\alpha/2}=1,96$ (dejando $\alpha/2=0,025$ a cada lado). Al ser $z < z_{\alpha/2}$ la hipótesis H_0 no se puede rechazar y por lo tanto no se puede aceptar que el nuevo producto sea mejor que el actual. El valor p obtenido, es decir la probabilidad de obtener valores del estadístico mayores que $0,19$ en el caso de que ambos productos sean iguales $P(|z| > 0,19) = 0,88$. Este valor p debe tenerse en cuenta en la decisión y no limitarse a ver si z supera a $z_{\alpha/2}$. En este caso, el valor p obtenido respalda la decisión de no rechazar la hipótesis H_0 .

Por otra parte, los valores de la desviación estándar de las muestras obtenidas son muy similares entre sí pero difieren bastante de la σ supuesta. Mediante el contraste¹ 3 de la Tabla 2 se puede comprobar si la diferencia es significativa. El planteamiento es:

- $H_0: \sigma = 0,35$.
- $H_1: \sigma \neq 0,35$.

$$\frac{ns^2}{\sigma^2} = \frac{10 * 0,19^2}{0,35^2} = 2,94$$

$$\chi^2_{10-1}(0,025) = 2,70 < \approx 2,94 \text{ (los datos rechazan "poco" que } \sigma < 0,35)$$

$$\chi^2_{10-1}(0,975) = 19,0 >>> 2,94 \text{ (los datos rechazan totalmente que } \sigma > 0,35)$$

2,94 es mayor que 2,70 pero está muy próximo a él, por lo que la hipótesis H_0 no está bien respaldada. Además, se puede hacer el contraste con el siguiente planteamiento:

- $H_0: \sigma = 0,35$.

¹ Este contraste requiere normalidad y es muy sensible a la falta de normalidad. Por esta razón se requiere comprobar la normalidad de los datos, que se hará más adelante.

HERRAMIENTAS ESTADÍSTICAS-COMPARACIÓN DE DOS MUESTRAS: t-test

- $H_1: \sigma < 0,35$.

Por los datos muestrales no se considera verosímil que $\sigma > 0,35$ y se prefiere concentrar el α a una cola. En este caso, el valor crítico del estadístico sería:

$$\chi^2_{10-1}(0,050) = 3,33 > 2,94$$

Por lo tanto debe rechazarse que $\sigma = 0,35$. Como consecuencia de esto, las conclusiones del contraste de igualdad de medias realizado, no tiene validez.

Los valores obtenidos de la varianza muestral son muy próximos, por lo que no parece necesario aplicar el contraste de 4 de la Tabla 2. Si se quisiera hacer con Minitab:

APLICACIÓN MINITAB

Minitab dispone de la opción **Stat->Basic Statistics->2Variances**. Eligiendo las opciones del cuadro de diálogo para comparar la igualdad de varianzas (H_0) frente a varianzas distintas (H_1), se obtiene la siguiente salida y la **¡Error! No se encuentra el origen de la referencia..**

7.3. COMPARACIÓN DE LA MEDIA DE DOS PROCESOS DE FABRICACIÓN. CASO VARIANZA DESCONOCIDA

EJEMPLO 14. COMPARACIÓN DE LA RESISTENCIA AL DESGASTE DE SUELAS DE ZAPATOS (CONTINUACIÓN)

Como consecuencia del análisis anterior, no puede aceptarse $\sigma = 0,35$. Si se analiza con el contraste 6 de la Tabla 2:

$$\frac{(\bar{x}_A - \bar{x}_B) - (\mu_A - \mu_B)}{\sqrt{\frac{(n_A - 1)s_A^2 + (n_B - 1)s_B^2}{n_A + n_B - 2} \left(\frac{1}{n_A} + \frac{1}{n_B} \right)}} = \frac{0,70 - 0,67}{\sqrt{\frac{(10 - 1)0,19^2 + (10 - 1)0,19^2}{10 + 10 - 2} \left(\frac{1}{10} + \frac{1}{10} \right)}} =$$

HERRAMIENTAS ESTADÍSTICAS-COMPARACIÓN DE DOS MUESTRAS: t-test

$$= \frac{0,70 - 0,67}{2 * 0,19 \sqrt{\frac{(10-1)}{10(10+10-2)}}} = 0,35$$

$$t_{10+10-2}(0,025) = 1,73 > 0,35$$

La hipótesis H_0 no se puede rechazar y por lo tanto no se puede aceptar que el nuevo producto sea mejor que el actual. El valor p es 0,70. El valor p obtenido respalda la decisión de no rechazar la hipótesis H_0 .

APLICACIÓN MINITAB

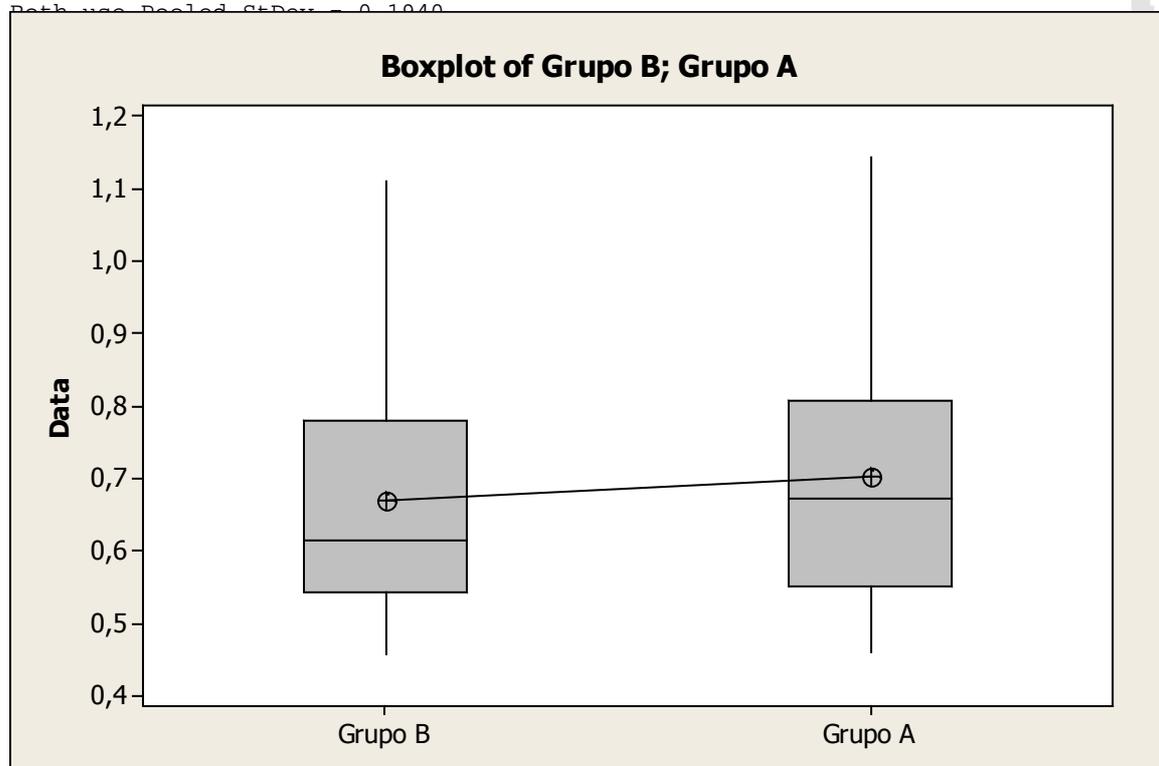
En Minitab esto se hace:

**Stat->Basic Statistics-> 2 Sample t
Two-Sample T-Test and CI: Grupo B; Grupo A**

Two-sample T for Grupo B vs Grupo A

	N	Mean	StDev	SE Mean
Grupo B	10	0,670	0,193	0,061
Grupo A	10	0,703	0,195	0,062

Difference = mu (Grupo B) - mu (Grupo A)
Estimate for difference: -0,033524
95% CI for difference: (-0,215811; 0,148763)
T-Test of difference = 0 (vs not =): T-Value = -0,39 P-Value = 0,704 DF = 18
Both use Pooled StDev = 0,1940



7.4. FORMACIÓN DE BLOQUES ALEATORIZADOS: CASO DE ENSAYOS EMPAREJADOS. VENTAJAS E INCONVENIENTES

En el caso anterior se compararon dos muestras, que aunque tenían el mismo tamaño, no existía ningún emparejamiento entre los datos. Esto presenta el inconveniente que la variabilidad existente dentro de cada muestra puede enmascarar la diferencia existente entre las medias de ambas. Por esta razón podría pensarse en un emparejamiento de los elementos de las muestras de manera que no difieran entre sí más que en el aspecto que se quiere comparar. Por ejemplo, si se está comparando un material de suela de zapatos, habría dos enfoques:

- Calzar 5 niños (10 pies) con los dos zapatos de un tipo y otros 5 niños con los zapatos de otro tipo. Esto es exactamente lo que se ha hecho en el caso anterior y se denomina **ensayo no pareado**. En este caso se mediría la diferencia de las medias de los desgastes de las suelas de cada grupo de chicos. Este modo de proceder tiene el inconveniente que mezcla el efecto “grupo” con el efecto “tipo de suela”. Si la conducta de los niños de ambos grupos fuera muy diferente podría llevar a una conclusión errónea.
- Calzar 10 niños con una suela de cada tipo, aleatorizando los pies derechos e izquierdos. En este caso se mediría la diferencia media del desgaste de los zapatos de cada chico. Este modo de proceder se denomina **ensayo pareado**. Obviamente, con este segundo modo de proceder se separan la diferencia del desgaste de otros factores. En este caso, la variable aleatoria d , a estudiar es también normal, al ser la resta de dos variables aleatorias que hemos supuesto normales.

HERRAMIENTAS ESTADÍSTICAS-COMPARACIÓN DE DOS MUESTRAS: t-test

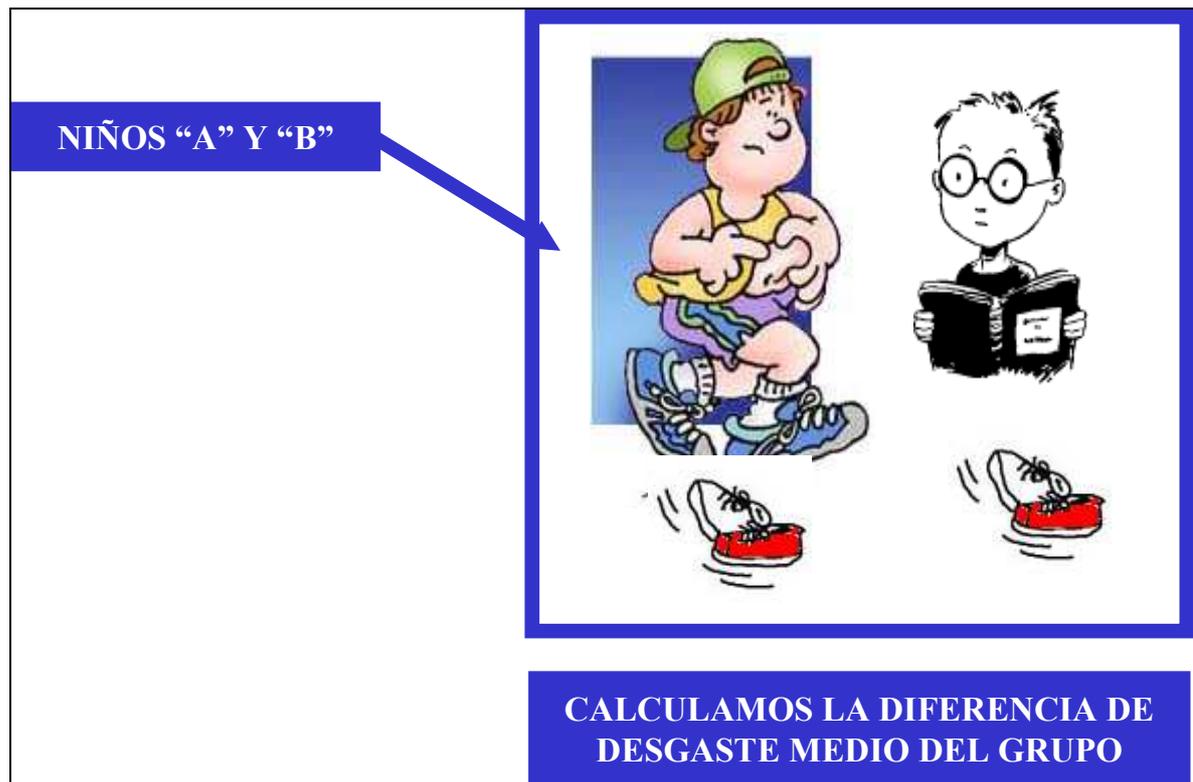


Figura 6: Ensayo pareado

El planteamiento es:

- $H_0: d = 0.$
- $H_1: d \neq 0.$

Como se trata de contrastar la media de una distribución normal con σ desconocida, el estadístico a emplear es el siguiente:

$$\frac{\bar{d}}{s_D / \sqrt{n}} \sim t_{n-1}$$

donde:

HERRAMIENTAS ESTADÍSTICAS-COMPARACIÓN DE DOS MUESTRAS: t-test

$$s_D^2 = \frac{\sum_{i=1}^n (d_i - \bar{d})^2}{n-1}$$

Si se considera la variable aleatoria “diferencia”, esta variable excluye el componente de la variabilidad debido a la heterogeneidad entre grupos, pero tenemos muchos menos grados de libertad. ¡En este caso se tienen únicamente 9 grados de libertad, mientras que en el anterior 18! Consecuentemente solo se gana sensibilidad en aquellos casos en los que la variabilidad debida a la heterogeneidad de la muestra sea relevante.

EJEMPLO 15. COMPARACIÓN DE LA RESISTENCIA AL DESGASTE DE SUELAS DE ZAPATOS. ENSAYO NO PAREADO

Supongamos que los mismos datos anteriores pero corresponden a los pies derechos e izquierdo de los diez niños.

Niño	Material A	Material B	d
1	0,46	0,46	0,00
2	0,62	0,71	-0,09
3	0,62	0,67	-0,05
4	0,77	0,51	0,26
5	1,11	0,8	0,31
6	0,47	0,68	-0,21
7	0,66	0,66	0,00
8	0,57	1,15	-0,58
9	0,61	0,57	0,04
10	0,82	0,84	-0,02
MEDIA	0,67	0,71	-0,03

Tabla 4: Datos de desgaste de los zapatos

Realizando el contraste:

HERRAMIENTAS ESTADÍSTICAS-COMPARACIÓN DE DOS MUESTRAS: t-test

$$\frac{\bar{d}}{s_D / \sqrt{n}} = \frac{-0,03}{0,24 / \sqrt{10}} = -0,43$$

$$t_{10-1}(0,025) = 2,26 > 0,43$$

La hipótesis H_0 no se puede rechazar. El valor p es 0,70. El valor p obtenido respalda la decisión de no rechazar la hipótesis H_0 .

APLICACIÓN MINITAB

En Minitab esto se hace:

Stat->Basic Statistics-> Paired t

Paired T-Test and CI: Grupo B; Grupo A

Paired T for Grupo B - Grupo A

	N	Mean	StDev	SE Mean
Grupo B	10	0,669822	0,192621	0,060912
Grupo A	10	0,703346	0,195395	0,061789
Difference	10	-0,033524	0,246404	0,077920

95% CI for mean difference: (-0,209790; 0,142743)

T-Test of mean difference = 0 (vs not = 0): T-Value = -0,43 P-Value = 0,677

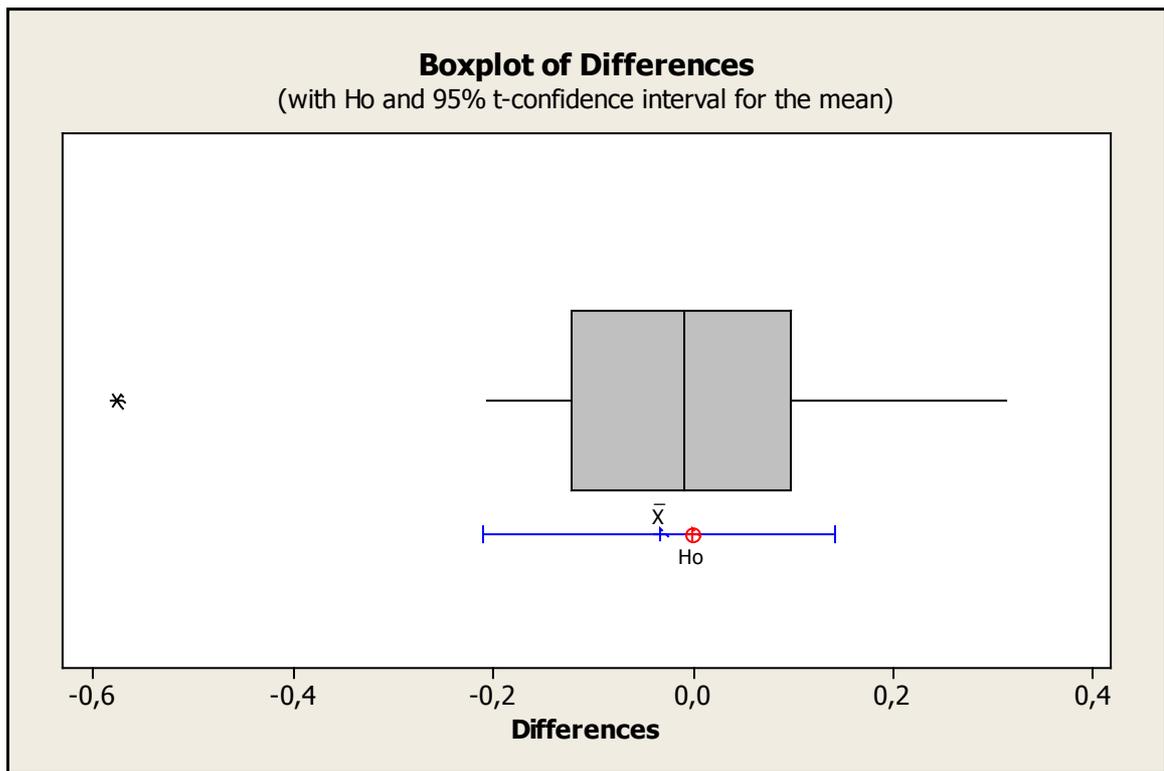


Figura 7: Diagrama de caja del desgaste en zapatos calzados por el mismo niño pero de distinto material

8. CONTRASTES DE AJUSTE

8.1. Concepto de contraste de ajuste

Una vez estimados los parámetros de la distribución a partir de los datos disponibles, es necesario proceder a validar el modelo de distribución estadística propuesta. Es decir, contrastar si la función de distribución propuesta se ajusta a los datos o si por el contrario las diferencias son tales que no permiten aceptar ese modelo como válido.

Los procedimientos más habituales para realizar estos contrastes de ajuste son los siguientes:

HERRAMIENTAS ESTADÍSTICAS-COMPARACIÓN DE DOS MUESTRAS: t-test

- χ^2 de Pearson (CS).
- Kolmogorov-Smirnov (KS).
- El método del papel probabilístico.
- Anderson-Darling (AD).

8.2. Contraste Chi-cuadrado de Pearson.

Es aplicable tanto a funciones de distribución discretas como continuas. En el caso de distribuciones continuas es necesario discretizarla previamente en intervalos, por lo que en lo sucesivo seguiremos el razonamiento suponiendo que la variable tiene carácter discreto.

Este contraste compara las frecuencias observadas y las frecuencias esperadas si la distribución propuesta es cierta. La operativa es la siguiente:

- Establecer la hipótesis H_0 de la función de probabilidad a emplear.
- Dividir en k "celdas" C_1, C_2, \dots, C_k de modo que cada clase contenga al menos tres observaciones y existan al menos cinco clases. Estas clases no tienen por que ser de igual tamaño necesariamente.
- Contabilizar las frecuencias de los datos disponibles f_1, f_2, \dots, f_k para cada uno de las celdas.
- Empleando los datos, estimar por máxima verosimilitud los parámetros requeridos por la función de probabilidad. Por ejemplo, si se tratara de una distribución normal, serían dos parámetros, μ y σ , si fuera una binomial sería un parámetros, p , etc.
- Sea $f(x)$ la función de probabilidad teórica. Entonces las frecuencias esperadas teóricas serían $nf(c_1), nf(c_2), \dots, nf(c_k)$.
- El estadístico:

$$\sum_{i=1}^k \frac{(f_i - nf(c_i))^2}{nf(c_i)} \sim \chi^2_{k-p-1}$$

HERRAMIENTAS ESTADÍSTICAS-COMPARACIÓN DE DOS MUESTRAS: t-test

se distribuye según una χ^2 de Pearson de $k-p-1$ grados de libertad, donde p es el número de parámetros estimados por máxima verosimilitud. Fijado un nivel de confianza α (generalmente $\alpha=0,05$) si:

$$\chi^2(\alpha)_{k-p-1} > \sum_{i=1}^k \frac{(f_i - nf(c_i))^2}{nf(c_i)}$$

La función R no puede rechazarse como modelo

$$\chi^2(\alpha)_{k-p-1} < \sum_{i=1}^m \frac{(f_i - nf(c_i))^2}{nf(c_i)}$$

La función R debe rechazarse como modelo

Donde $\chi^2(\alpha)_{k-p-1}$ se denomina “nivel crítico para el intervalo de confianza α ”.

Obsérvese una vez más que, de acuerdo con la metodología clásica del contraste estadístico de hipótesis, en ningún caso se llega a “aceptar” sino que simplemente “no hay evidencia para rechazarla”. En general la decisión no debe ser ciega en base al valor obtenido en la ecuación anterior, sino que debe hallarse el nivel de confianza que correspondería a ese valor como nivel crítico.

EJEMPLO 16. ¿SE DISTRIBUYEN LOS DEFECTOS SEGÚN UNA DISTRIBUCIÓN DE POISSON?

Se está estudiando cómo se distribuyen los defectos en unas piezas de fabricación y se quiere contrastar si los defectos se distribuyen según la función de Poisson. Esto tiene su importancia, porque si los defectos siguen una distribución de Poisson, una pieza que tenga ya un defecto tiene las mismas probabilidades de tener otro defecto que cualquier otra pieza.

La metodología es la siguiente:

HERRAMIENTAS ESTADÍSTICAS-COMPARACIÓN DE DOS MUESTRAS: t-test

Núm. Defectos	frecuencia	Número de defectos	POISSON	Frec. Esperada	(fre -fre esp)2/fre esp
0	229	0	0,395	227,53	0,01
1	211	211	0,367	211,34	0,00
2	93	186	0,170	98,15	0,27
3	35	105	0,053	30,39	0,70
4	7	28	0,012	7,06	0,00
5	1	5	0,002	1,31	0,07
TOTAL	576	535	1,000	575,77	1,05

Defectos / pieza 0,929

En este caso, $K = 6$, $p = 1$, $\chi^2(0,05)_{6-1} = 9,49 \gg 1,05$. por lo que no se puede rechazar la función de Poisson.

EJEMPLO 17. TABLAS DE CONTINGENCIA

Se está estudiando la influencia de la duración de un tratamiento térmico en la fragilidad de unas piezas. Se obtuvieron los siguientes resultados.

Duración del tratamiento (h)	Número de piezas frágiles	Número de piezas resilientes	Total
30	77	323	400
60	177	223	400
Total	254	546	800

La hipótesis de indiferencia del tratamiento en que la pieza sea frágil o resiliente es equivalente a suponer que la función de distribución es una binomial de $p = 400 / 400$. En general se distribuye según una $\chi^2_{(r-1)(c-1)}$ donde r y c son el número de filas y columnas. Minitab tiene un procedimiento para resolverlo directamente:

APLICACIÓN MINITAB

Stat-> Tables-> Cross tabulation and Chi-Square

Se obtiene:

Tabulated statistics: Resultado; Longitud de ciclo

Using frequencies in Cantidad

HERRAMIENTAS ESTADÍSTICAS-COMPARACIÓN DE DOS MUESTRAS: t-test

Rows: Resultado Columns: Longitud de ciclo

	30	90	All
Frágil	77	177	254
	127	127	254
	19,685	19,685	*
Resiliente	323	223	546
	273	273	546
	9,158	9,158	*
All	400	400	800
	400	400	800
	*	*	*

Cell Contents: Count
Expected count
Contribution to Chi-square

Pearson Chi-Square = 57,685; DF = 1; P-Value = 0,000
Likelihood Ratio Chi-Square = 58,875; DF = 1; P-Value = 0,000

Como se puede ver la hipótesis de que la duración del tratamiento no influye no se puede mantener.

8.3. Contraste de Kolmogorov Smirnov (KS)

Se aplica únicamente a variables continuas. Si los datos obtenidos son:

$$X_1 \leq X_2 \leq \dots X_n,$$

La función de distribución empírica de la muestra, $F_n(x)$ se calcula:

$$F_n(x) = \begin{cases} 0 & x < x_1 \\ \frac{r}{n} & x_r \leq x \leq x_{r+1} \\ 1 & x \geq x_n \end{cases}$$

A continuación, se construye el estadístico que marca la máxima distancia entre los valores dados por la función empírica y los teóricos de la función que se está contrastando el ajuste:

$$D_n = \max |F_n(x) - F(x)|$$

Este estadístico está tabulado.

8.4. Contraste de Anderson-Darling (AD)

El contraste de Anderson-Darling es una modificación del KS que utiliza el siguiente estadístico que proporciona más peso a los valores de las colas.

$A = N - S$, donde S es:

$$S = \sum_{i=1}^N \frac{(2i-1)}{N} [\ln F(Y_i) + \ln(1 - F(Y_{N+1-i}))]$$

8.5. Aplicación a los contrastes de normalidad

Dado que la hipótesis de normalidad se requiere para la mayor parte de las aplicaciones que se han visto, resulta muy importante comprobar si los datos disponibles son consistentes con la distribución normal. Para ellos puede utilizarse los contrastes anteriores (de los que existen publicadas variantes específicas para contrastar normalidad).

Para muestras pequeñas resulta bastante práctico el uso del papel probabilístico. Este papel está construido de tal forma que, si se representan los puntos $(x, F(x))$, éstos quedan en una línea recta si son datos normales. Este test también se conoce como “*recta de Henry*”.

HERRAMIENTAS ESTADÍSTICAS-COMPARACIÓN DE DOS MUESTRAS: t-test

Papel Probabilístico: Datos normales

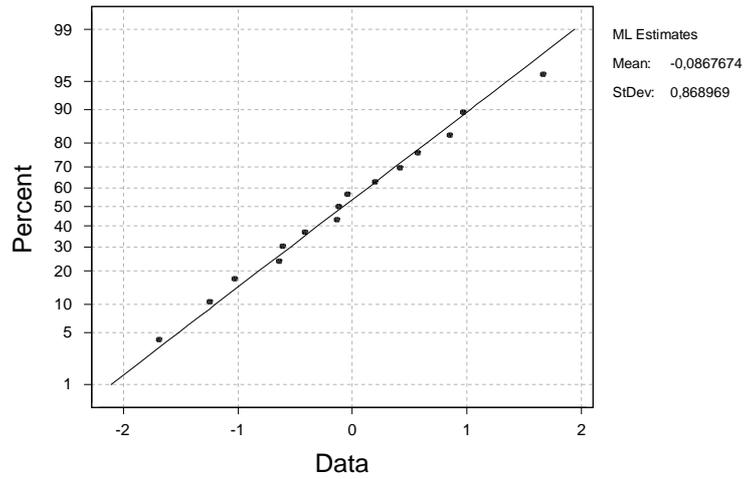


Figura 8: Papel probabilístico normal. Datos normales

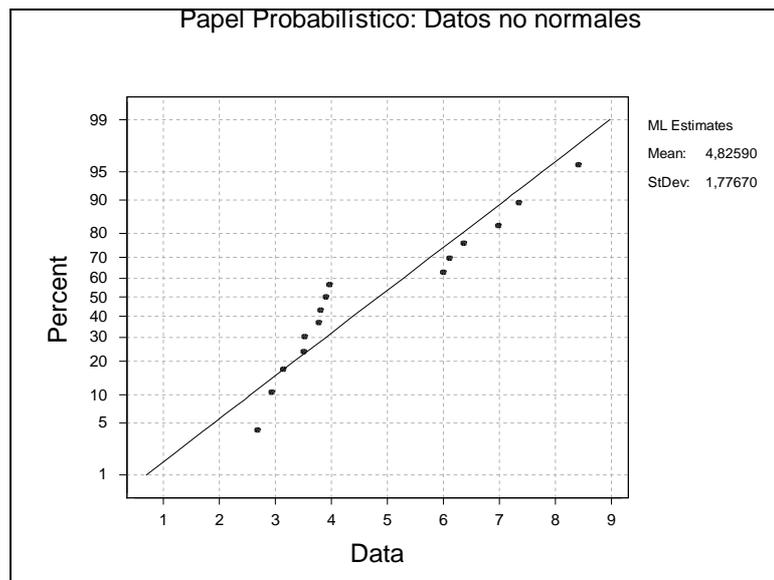


Figura 9: Papel probabilístico normal. Datos no normales

EJEMPLO 18. COMPROBAR QUE LOS DATOS DEL EJEMPLO 13 SON NORMALES

APLICACIÓN MINITAB

Hay dos procedimientos

Stat-> Basic Statistics-> Normality test

Ver Figura 10.

Graf-> Probability Plot-> Normal

Ver Figura 11.

De estos tests se deduce que la normalidad no se puede rechazar.

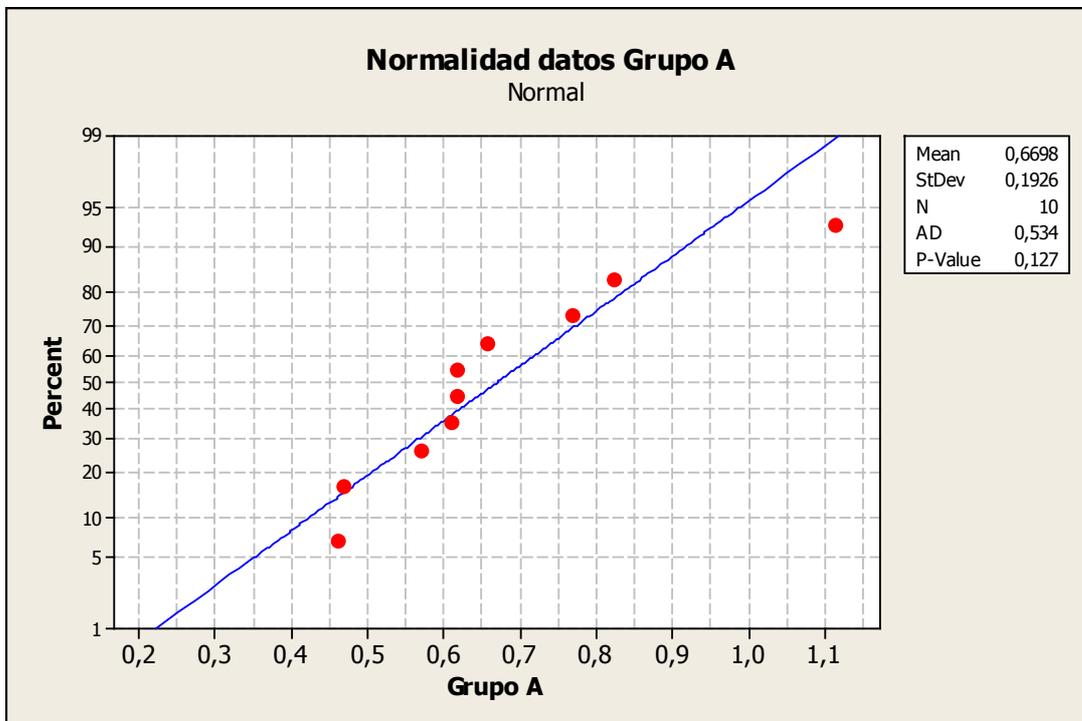


Figura 10: Test de normalidad datos de los zapatos de los niños grupo A

HERRAMIENTAS ESTADÍSTICAS-COMPARACIÓN DE DOS MUESTRAS: t-test

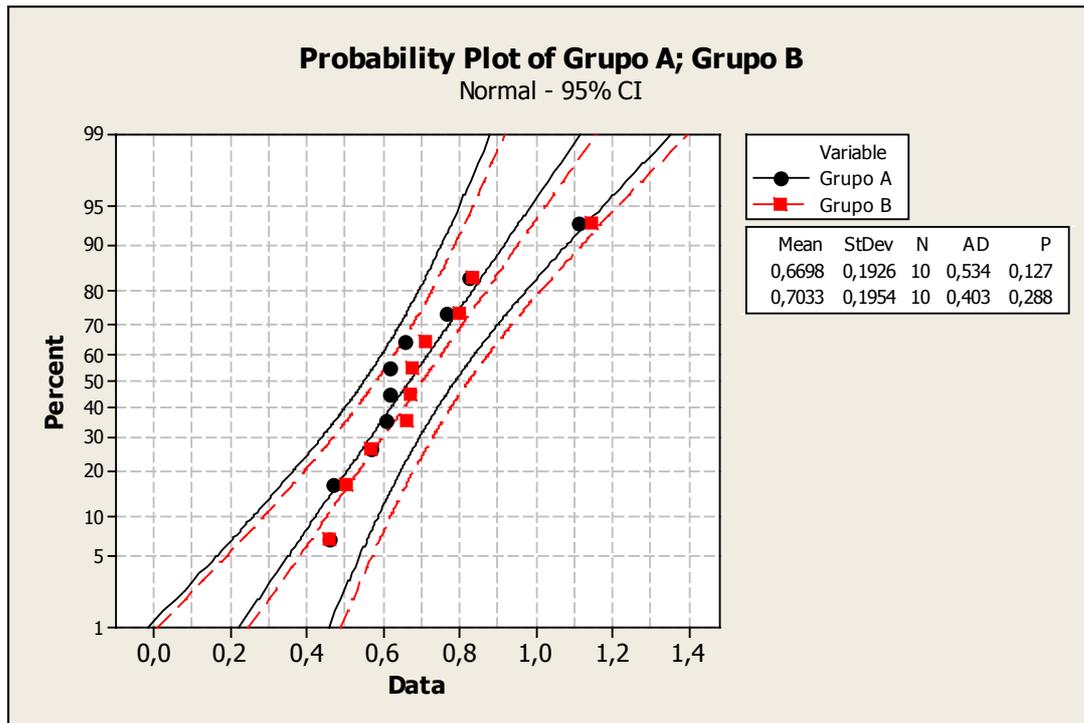


Figura 11: Test de normalidad superpuesto de los datos de los zapatos de los niños de los grupos A y B

9. BATERÍA DE EJERCICIOS DE CONTRASTES

EJEMPLO 19. CONTRASTE DE LA MEDIA

Los datos de la tabla siguiente proceden de una muestra de piezas. Se pretende contrastar la hipótesis de que el proceso se encontraba ajustado a 51.

52.8722	51.5258	51.2244	51.5442	50.9489	49.9980	50.0832
51.2512	49.6413	50.1273	50.6276	52.1107	50.9617	50.6465
50.2696	51.3255	52.3583	49.3578	51.1400	50.3389	52.8284
50.9583	49.7253	49.6357	49.7463	49.8839	50.7907	49.1591
51.6626	49.1528					

Tabla 5

SOLUCIÓN:

HERRAMIENTAS ESTADÍSTICAS-COMPARACIÓN DE DOS MUESTRAS: t-test

$$H_0 : \mu = 51$$

$$H_1 : \mu \neq 51$$

Aplicando el test t:

$$\bar{x} = 50,73 \quad s = 1,009 \quad n = 30 \quad t_{29}(0,025) = 2,045$$

$$\sqrt{29} \frac{50,73 - 51}{1,009} = -1,44$$

1,44 < 2,045 luego no se puede rechazar la hipótesis con un nivel de significación del 95%.

EJEMPLO 20. EJERCICIO 3

Se extrae una segunda muestra del proceso descrito en los ejercicios anteriores y cuyos resultados se recogen en la tabla siguiente. ¿Proporcionan estos datos la evidencia de que el proceso se ha desajustado?

50.5239	50.0935	51.3964	50.4915	52.7089	50.0232	51.1062
51.2792	51.2210	51.1323	52.0471	50.6068	50.7610	51.6288
50.8009	49.7507	49.7470	49.0381	50.3587	51.3914	52.8860
50.7293	50.0995	52.0671	52.0722	50.7513	51.1660	51.3476
51.3262	51.5460					

Tabla 6

SOLUCIÓN:

$$\bar{x}_1 = 50,73 \quad \bar{x}_2 = 51,00 \quad s_1 = 1,009 \quad s_2 = 0,858 \quad n = 30$$

$$t_{58}(0,025) = 2 \quad \mu = 51$$

HERRAMIENTAS ESTADÍSTICAS-COMPARACIÓN DE DOS MUESTRAS: t-test

$$s_t^2 = \frac{30(1,009^2 + 0,858^2)}{30 + 30 - 2} = 0,9074$$

$$\frac{50,73 - 51,00}{\sqrt{0,9074} \sqrt{1/30 + 1/30}} = -1,1099$$

1,1099 < 2 luego no se puede rechazar la hipótesis con un nivel de significación del 95%.

EJEMPLO 21. ¿SE DISTRIBUYE NORMALMENTE?

Compruebe si se distribuye el proceso normalmente.

Utilizando los gráficos de papel probabilístico:

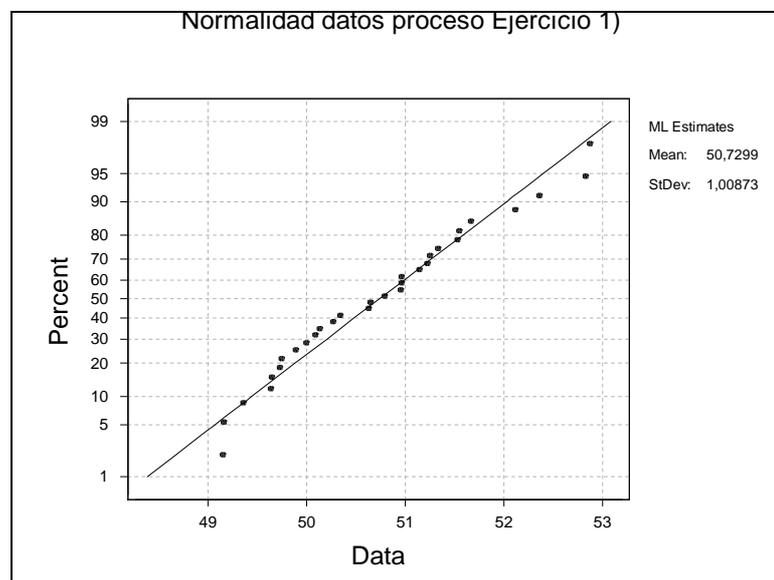


Fig. 1 Datos del proceso inicial

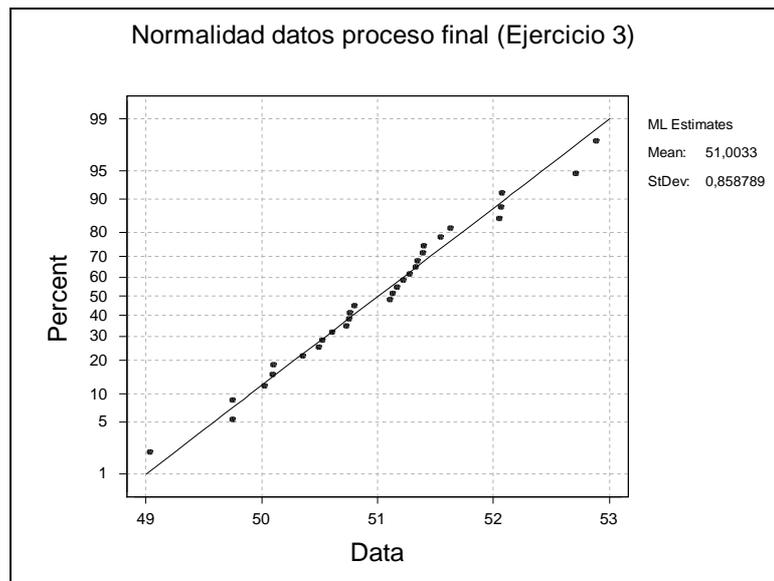


Fig. 2 Datos del proceso final

9.1. CASO NOVOMAT

9.1.1. Escena 1

NOVOMAT, S.A. es una empresa dedicada a la fabricación de piezas en materiales compuestos a base de fibra de vidrio, kevlar y fibra de carbono. Sus clientes son empresas de alta tecnología que precisan incorporar a sus diseños componentes con un elevado límite elástico y bajo peso. Las exigencias de estos clientes son cada vez mayores y las especificaciones de compra cada vez más complejas.

Para mejorar el proceso, NOVOMAT está evaluando dos alternativas, la "A", que es la actualmente implantada y la "B" que incorpora modificaciones sustanciales al proceso. Los ingenieros de proceso basándose en datos anteriores saben que:

- La distribución de la característica de calidad en las piezas es sensiblemente normal.
- La desviación típica del proceso es sensiblemente constante igual a 1,5.

HERRAMIENTAS ESTADÍSTICAS-COMPARACIÓN DE DOS MUESTRAS: t-test

Para comparar las dos alternativas se fabrican dos muestras de 10 unidades. Una vez ensayadas en el laboratorio, se obtienen los siguientes valores:

A	10.57	12.08	8.84	10.44	10.36	10.17	9.94	10.26	10.77	10.01
B	11.19	11.07	10.10	11.01	11.84	10.43	9.89	11.10	11.24	10.64

Tabla 7: Valores de ensayo de las muestras de las alternativas A y B.

9.1.2. Escena 2

NOVOMAT, S.A. considera que no cuenta con la suficiente seguridad para suponer que la desviación típica del proceso es igual a 1,5 y desea repetir el análisis sin esa suposición.

9.1.3. Escena 3

NOVOMAT está comparando dos procesos de pintado. Para eliminar el efecto de la heterogeneidad de las placas, se ha realizado un ensayo pareado, es decir que la mitad de una placa se pintaría con un proceso y la otra mitad con otro. Se supone que no influye el hecho de que cuando se pinte sobre parte de la placa (por ejemplo una cara), ya se haya pintado sobre otra parte de la placa (por ejemplo la otra cara). Los resultados obtenidos son los de la Tabla 8. Obsérvese que coinciden numéricamente con los del ejemplo anterior.

A	10.57	12.08	8.84	10.44	10.36	10.17	9.94	10.26	10.77	10.01
B	11.19	11.07	10.10	11.01	11.84	10.43	9.89	11.10	11.24	10.64
Delta	-0.62	1.01	-1.26	-0.57	-1.48	-0.26	0.05	-0.84	-0.47	-0.63

Tabla 8: Datos del ensayo (pareado)