**PRUEBA T PARA UNA MUESTRA**

El procedimiento Prueba T para una muestra contrasta si la media de una sola variable difiere de una constate especificada. Es decir, se emplea en situaciones en las que se pretende contrastar la diferencia entre la media de una muestra y la de su población de origen. Así podemos saber si una determinada muestra procede de una población cuya media verdadera se conoce. También este procedimiento puede ser muy útil. Por ejemplo, un ingeniero desea comprobar si la puntuación media del

coeficiente intelectual de un grupo de alumnos difiere de 100 con un nivel de confianza al 95%.

**PROCEDIMIENTO PARA OBTENER UNA PRUEBA T PARA UNA MUESTRA**

* Realizar el Análisis exploratorio y descriptivo.
* Verificar el supuesto de normalidad.
* Realizar el contraste para la media

1. **PRUEBA T PARA UNA MUESTRA: A PARTIR DE UNA POBLACIÓN CON DISTRIBUCIÓN NORMAL Y VARIANZA DESCONOCIDA**

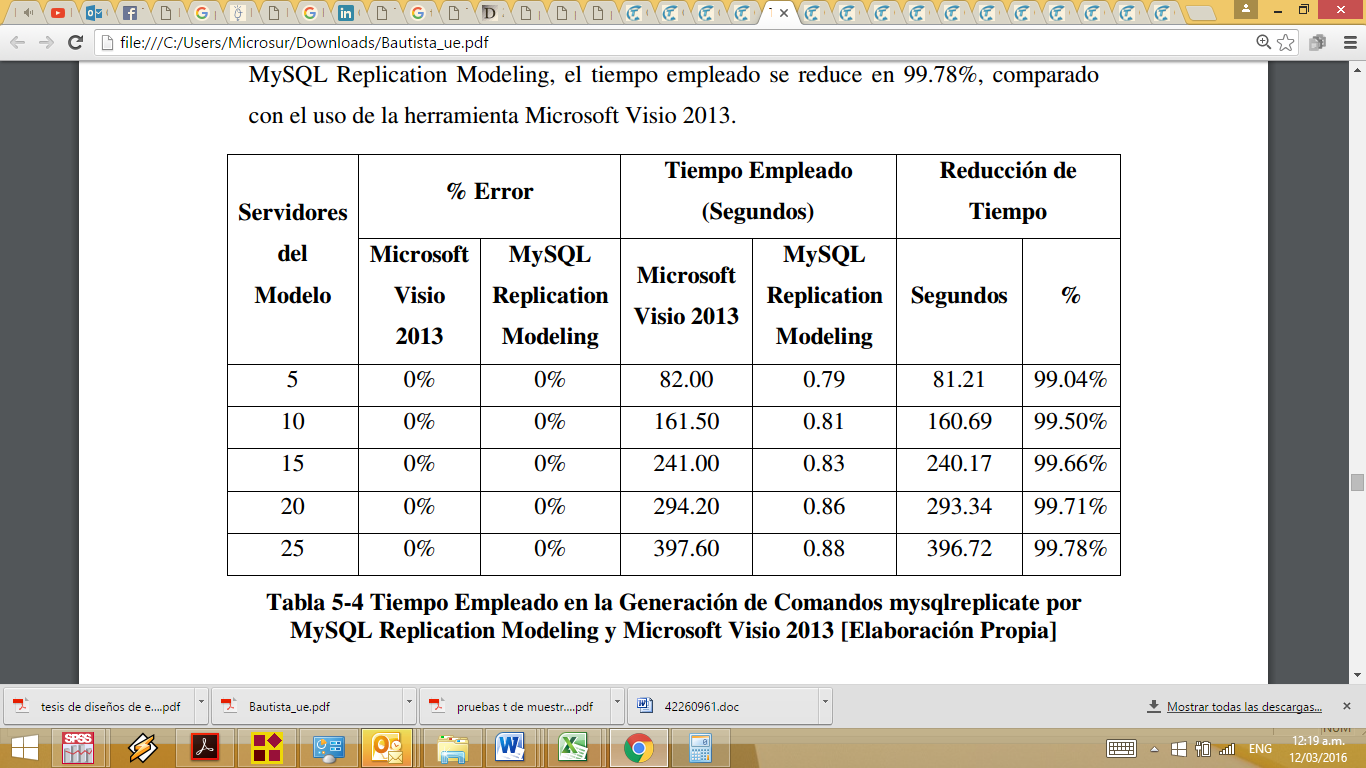
Tesis: Herramienta para el modelado de la replicación de MySql basada en la ingeniería dirigida por modelos

Autor: Efraín Ricardo Bautista Ubillús

Ejemplo: La tesis tiene como objetivo generar automáticamente a partir de un modelo válido, los comandos mysqlreplicate de configuración, y reducir en más del 99% el tiempo empleado usando el método manual con Microsoft Visio 2013.

Solución:

1. Datos: Tiempo Empleado en la Generación de Comandos mysqlreplicate por MySQL Replication Modeling y Microsoft Visio 2013



|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Reducción de tiempo (%)** | 99.04 | 99.50 | 99.66 | 99.71 | 99.78 |

1. Supuestos: Las cinco (5) muestras se toman de manera aleatoria de una población de individuos con las mismas características.
2. Enunciado de Hipótesis

Ho: Los comandos mysqlreplicate de configuración no reducen en más del 99% el tiempo empleado usando el método manual con Microsoft Visio 2013.

H1: Los comandos mysqlreplicate de configuración reducen en más del 99% el tiempo empleado usando el método manual con Microsoft Visio 2013.

1. Hipótesis estadística

Ho: µ ≤ 99.0%

H1: µ > 99.0%

1. Realizar la prueba de Normalidad.

Según la prueba de Normalidad realizada en SPSS, ambas poblaciones, provienen de una distribución normal.

Ho: Los datos provienen de una población normal

H1: Los datos no provienen de una población normal



Dado que el valor – p es mayor que el nivel de significancia (0.200), entonces no se rechaza la H0, por lo tanto los datos provienen de una población normal.

1. Estadística de Prueba: Dado que se desconoce la varianza de la población, la estadística de prueba se obtiene mediante la ecuación.

C:\Documents and Settings\qwer\Mis documentos\Mis imágenes\f338ec74d591e1bcf826d953fa21b1e0.png

1. Distribución de la estadística de prueba: Sigue una distribución t de Student, con n-1=5-1=4 grados de libertad.
2. Regla de Decisión: Sea α=0.05. Dado que se tiene una prueba unilateral, se pone α=0.05 en una cola de la distribución de la estadística de prueba. Los valores de Tt (T de tabla) a la derecha el área está definida hasta +2.13

C:\Documents and Settings\qwer\Mis documentos\Mis imágenes\f338ec74d591e1bcf826d953fa21b1e0.png

1. Cálculo de la estadística de prueba. = 99.54 – 99.0 / 0.133 = Tc = +4.06
2. Decisión estadística: Se rechaza Ho, dado que +4.06 cae en la región de rechazo.
3. Conclusión: La conclusión, con base a estos datos, es que a un margen de error del 5%, se concluye que los comandos mysqlreplicate de configuración reducen en más del 99% el tiempo empleado usando el método manual con Microsoft Visio 2013.

A un margen de error del 5% se concluye, que los comandos mysqlreplicate de configuración reducen en más del 99% el tiempo empleado usando el método manual con Microsoft Visio 2013.

**Ejemplo:** La especialista en recursos humanos de una importante corporación está reclutando un gran número de empleados para un proyecto en el extranjero. Durante el proceso de selección, la administración le pregunta cómo van las cosas, y ella responde: “Bien. Creo que la puntuación promedio en la prueba de aptitudes será aproximadamente 90”. Se pretende conocer si es posible concluir que el promedio de las puntuaciones para la población de la que se extrajo la muestra no es 90.

Solución:

1. Datos

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 96 | 95 | 72 | 98 | 89 | 77 | 83 | 96 | 95 | 64 |
| 98 | 79 | 89 | 83 | 69 | 94 | 78 | 74 | 67 | 92 |

1. Supuestos: Los 20 individuos constituyen una muestra aleatoria de una población de individuos con las mismas características
2. Hipótesis

Ho: µ = 90

H1: µ ≠ 90

1. Realizar la prueba de Normalidad.

Según la prueba de Normalidad realizada en SPSS, ambas poblaciones, provienen de una distribución normal.

Ho: Los datos provienen de poblaciones normales

H1: Los datos no provienen de poblaciones normales

1. Estadística de Prueba: Dado que se desconoce la varianza de la población, la estadística de prueba se obtiene mediante la ecuación.

C:\Documents and Settings\qwer\Mis documentos\Mis imágenes\f338ec74d591e1bcf826d953fa21b1e0.png

1. Distribución de la estadística de prueba: Sigue una distribución t de Student, con n-1=20-1=19 grados de libertad.
2. Regla de Decisión: Sea α=0.05. Dado que se tiene una prueba bilateral, se pone α/2=0.025 en cada cola de la distribución de la estadística de prueba. Los valores de t a la derecha e izquierda de los cuales están 0.025 del área son -2.093 y 2.093

C:\Documents and Settings\qwer\Mis documentos\Mis imágenes\f338ec74d591e1bcf826d953fa21b1e0.png

1. Cálculo de la estadística de prueba. = (-5.6) / 2.52 = -2.227
2. Decisión estadística: Se rechaza Ho, ya que -2.227 cae en la región de rechazo.
3. Conclusión: Con base a estos datos, es que a un nivel de significancia del 0.05, el promedio de las puntuaciones para la población de la que se extrajo la muestra no es igual a 90.

**PRUEBA T DE STUDENT PARA DOS MUESTRAS INDEPENDIENTES**

El procedimiento Prueba T para dos muestras independientes (muestras no relacionadas) compara las medias de dos grupos de casos.

Se usa para diseños experimentales en los cuales se estudia una variable independiente bajo dos condiciones.

Para esta prueba, idealmente los sujetos se deben de asignar aleatoriamente a dos grupos, de forma que cualquier diferencia en las respuestas sea debido al tratamiento y no a otros factores.

Este procedimiento permite comparar las medias de dos grupos diferentes para sacar conclusiones sobre lo que no les es común.

Para obtener una prueba T para muestras independientes, se realiza el siguiente procedimiento:

Las muestras de los grupos investigados no necesariamente deben tener el mismo tamaño.

Se deben obtener las medias separadamente de cada grupo.

**PROCEDIMIENTO PARA OBTENER UNA PRUEBA T PARA MUESTRAS INDEPENDIENTES**

* Realizar el Análisis exploratorio y descriptivo.
* Verificación de supuestos de normalidad y de homogeneidad.
* Realizar la prueba T de comparación de muestras independientes.

1. **PRUEBA T PARA DOS MUESTRAS INDEPENDIENTES QUE PRESENTA UNA DISTRIBUCIÓN NORMAL**

Tesis: Herramienta para el modelado de la replicación de MySql basada en la ingeniería dirigida por modelos

Autor: Efraín Ricardo Bautista Ubillús

A nivel de modelado, diversas herramientas de diagramación, tales como Microsoft Visio entre otras, son usadas para diseñar modelos de replicación MySQL. Sin embargo, este tipo de herramientas no permiten validar automáticamente si un modelo de replicación de MySQL está libre de errores, lo que puede resultar tener documentación errónea de los modelos de replicación. Dichas herramientas tampoco permiten generar automáticamente a partir de un modelo de replicación MySQL los comandos mysqlreplicate de configuración. La falta de la funcionalidad antes mencionada conlleva a realizarlas de forma manual, la cual es una tarea tediosa, propensa a errores y que consume tiempo, más aún si el número de servidores MySQL del modelo es alto, con 15, 20, 25 o más servidores. ¿Existen diferencias entre el tiempo empleado (segundos) en la Corrección de un Modelo de Replicación MySQL por MySQL Replication Modeling y Microsoft Visio 2013? ? Establezca y pruebe la hipótesis adecuada para un α=0.05

Solución:

1. Datos

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Microsft Visio 2013** | 92.20 | 193.00 | 331.00 | 421.10 | 595.70 |
| **MySQL Replication Modeling** | 34.40 | 49.30 | 51.30 | 53.70 | 74.90 |

1. Enunciado de hipótesis

Ho: No existen diferencias entre el tiempo empleado (segundos) en la Corrección de un Modelo de Replicación MySQL por MySQL Replication Modeling y Microsoft Visio 2013

H1: Existen diferencias entre el tiempo empleado (segundos) en la Corrección de un Modelo de Replicación MySQL por MySQL Replication Modeling y Microsoft Visio 2013

1. Hipótesis

Ho: µA  = µB

H1: µA  ≠ µB

1. Nivel de significación: α = 0.05
2. Realizar el análisis exploratorio de datos y descriptivo



Muestra 1 (Microsoft Visio 2013): s1 = 196.31; n1 = 5; http://www.hiru.com/matematika/matematika_06400.html/matematicas_064_03p.gif1=326.60

Muestra 2: (por MySQL Replication Modeling): s2 = 14.51; n2 = 5; http://www.hiru.com/matematika/matematika_06400.html/matematicas_064_03p.gif2=52.72

1. Prueba de Normalidad. Según la prueba de Normalidad realizada en SPSS, ambas poblaciones, provienen de una distribución normal.

Ho: Los datos provienen de poblaciones normales

H1: Los datos no provienen de poblaciones normales



Dado que el valor – p es mayor que el nivel de significancia (0.200), entonces no se rechaza la H0, por lo tanto los datos provienen de poblaciones normales, a un nivel de significancia del 0.05.

1. Prueba de Homogeneidad de varianzas: Según la prueba de Homogeneidad de varianzas (Prueba de Levene) realizada en SPSS, Se han asumido varianzas iguales.

Ho: Se han asumido varianzas iguales

H1: No se han asumido varianzas iguales



Dado que el valor – p es menor que el nivel de significancia (0.200), entonces se rechaza la H0, por lo tanto, No se han asumido varianzas iguales, a un nivel de significancia del 0.05.

1. Estadística de Prueba: Para realizar la prueba de hipótesis utilizaremos en este caso la prueba “t”. El estadístico “t”, queda establecido como Prueba “t” para dos muestras independientes, con distribución normal, las varianzas poblacionales no se conocen y se han asumido varianzas diferentes. Grados de libertad 17,362.

****

1. Decisión: Como el valor-p (0.000) < 0.05, rechazamos Ho.
2. Conclusión: A un nivel de significancia del 0.05, se concluye que, existen diferencias entre el tiempo empleado (segundos) en la Corrección de un Modelo de Replicación MySQL por MySQL Replication Modeling y Microsoft Visio 2013

A un margen de error del 5%, y dado que las varianzas son diferentes se concluye que existen diferencias entre el tiempo empleado (segundos) en la Corrección de un Modelo de Replicación MySQL por MySQL Replication Modeling y Microsoft Visio 2013

1. **PRUEBA T PARA DOS MUESTRAS INDEPENDIENTES QUE NO PRESENTA UNA DISTRIBUCIÓN NORMAL**

Un educador deseaba saber si era posible concluir que los estudiantes de primer año, procedentes de escuelas rurales y de escuelas urbanas son diferentes respecto de los puntajes de razonamiento verbal promedio. Una muestra aleatoria de 15 estudiantes venidos de una escuela rural y 18 de una escuela urbana obtuvieron los promedios que se muestran en la siguiente tabla en una prueba de razonamiento verbal. Use el nivel de significación del 5%

Solución:

1. Datos

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Estudiantes de escuelas rurales | 66 | 73 | 63 | 65 | 80 | 74 | 69 | 80 | 63 | 54 | 42 | 80 | 57 | 79 | 52 |  |  |  |
| Estudiantes de escuelas urbanas | 68 | 91 | 89 | 60 | 73 | 90 | 79 | 97 | 64 | 62 | 91 | 69 | 64 | 89 | 99 | 76 | 77 | 96 |

1. Hipótesis

Ho: Los estudiantes del primer año, procedentes de escuelas rurales y de escuelas urbanas son iguales respecto de los puntajes de razonamiento verbal promedio.

H1: Los estudiantes del primer año, procedentes de escuelas rurales y de escuelas urbanas son iguales respecto de los puntajes de razonamiento verbal promedio.

1. Nivel de significación: α = 0.05
2. Realizar el análisis exploratorio de datos y descriptivo

****

Muestra 1(rurales): sA = 11.624; nA = 15; http://www.hiru.com/matematika/matematika_06400.html/matematicas_064_03p.gifA=66.47

Muestra 2: (urbanas): sB = 13.249; nB = 18; http://www.hiru.com/matematika/matematika_06400.html/matematicas_064_03p.gifB=79.67

1. Prueba de Normalidad. Según la prueba de Normalidad realizada en SPSS. Los datos no provienen de poblaciones normales.

Ho: Los datos provienen de poblaciones normales

H1: Los datos no provienen de poblaciones normales

****

1. Dado que no se cumple el supuesto de normalidad la prueba a utilizar para comparar los dos grupos es la prueba de Mann Whitney (prueba no paramétrica).
2. Estadística de Prueba No Paramétrica, para dos muestras independientes: Prueba de Mann Whitney.

****

1. Decisión: Como p<0.05 rechazamos Ho.

**PRUEBA T DE STUDENT PARA DATOS PAREADOS O MUESTRAS RELACIONADAS**

El procedimiento Prueba T para muestras relacionadas compara las medias de un mismo grupo y calcula las diferencias entre la primera y la segunda medición. Se suele conocer como pruebas “Ex – Post” (Antes y Después)

Por ejemplo el efecto de un programa de capacitación a un grupo de la empresa. Se toma una prueba antes del entrenamiento y la misma después del entrenamiento.

**PROCEDIMIENTO PARA OBTENER UNA PRUEBA T PARA MUESTRAS PAREADAS O MUESTRAS RELACIONADAS**

* Ingresar los datos.
* Realizar el Análisis exploratorio y descriptivo.
* Verificación de supuestos de normalidad de las diferencias.
* Prueba de comparación de las medias.

1. **PRUEBA T PARA DOS MUESTRAS RELACIONADAS QUE PRESENTA UNA DISTRIBUCIÓN NORMAL**

Ejemplo: En un estudio acerca del estado general de la salud mental de personas retiradas, un equipo de psicólogos seleccionó una muestra de aquellas que deberían retirarse dentro de los próximos dos años.

Los investigadores desean saber si se podía sacar como conclusión que el nivel de salud mental de la población de la cual se había sacado la muestra tendía a tener diferencias antes y después del retiro. Use un niel de significancia del 5%.

Solución:

1. Datos

Estado de salud mental de 15 sujetos antes y después del retiro

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| SUJETO | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| Antes | 76 | 80 | 86 | 87 | 85 | 95 | 97 | 75 | 87 | 96 | 98 | 77 | 80 | 87 | 89 |
| Después | 70 | 75 | 84 | 90 | 81 | 95 | 87 | 72 | 92 | 85 | 88 | 76 | 85 | 81 | 84 |

1. Hipótesis

Ho: El nivel de salud mental de la población de la cual se había sacado la muestra no tendía a tener diferencias antes y después del retiro.

H1: El nivel de salud mental de la población de la cual se había sacado la muestra tendía a tener diferencias antes y después del retiro.

1. Nivel de significación: α = 0.05
2. Realizar el análisis exploratorio de datos y determinación de las medidas descriptivas de las variables antes y después.



1. Generar la variable diferencias ***(di) = Antes - Después***
2. Verificación de Supuestos de Normalidad para las diferencias (di)
   * Determinar las diferencias ***(di)***
   * Realizar la prueba de normalidad para las diferencias ***(di)***

Ho: Los datos de las diferencias provienen de poblaciones normales

H1: Los datos de las diferencias no provienen de poblaciones normales

****

1. Estadística de Prueba: Para realizar la prueba de hipótesis utilizaremos en este caso la prueba “t”. El estadístico “t”, queda establecido como Prueba “t” para dos muestras relacionadas, con distribución normal.



1. Verificación de la hipótesis de investigación.
2. Decisión: Como p<0.05, rechazamos Ho.

El nivel de salud mental de la población de la cual se había sacado la muestra tendía a tener diferencias antes y después del retiro.

1. **PRUEBA T PARA DOS MUESTRAS RELACIONADAS QUE NO PRESENTA UNA DISTRIBUCIÓN NORMAL**

Ejemplo: Veinte estudiantes del quinto año de Educ. Secundaria participaron en un experimento preparado para evaluar la efectividad de un programa para aumentar la comprensión de lectura. Cada sujeto hizo una prueba de comprensión de lectura antes y después de participar en el programa. ¿Proporcionan estos datos suficiente evidencia como para indicar que el programa es efectivo a un nivel de significancia del 0.05?

Solución:

1. Datos

Puntaje en comprensión de lectura de 20 estudiantes de quinto año de Educación Secundaria antes y después de participar en un programa preparado para mejorar la comprensión de lectura.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| SUJETO | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| Antes | 80 | 65 | 73 | 60 | 60 | 73 | 76 | 53 | 57 | 65 | 61 | 54 | 77 | 66 | 59 | 74 | 52 | 56 | 63 | 67 |
| Después | 85 | 75 | 84 | 55 | 72 | 81 | 84 | 50 | 65 | 75 | 79 | 62 | 85 | 78 | 69 | 88 | 60 | 52 | 75 | 80 |

1. Hipótesis

Ho: No existe diferencia entre los puntajes de comprensión de lectura antes y después de participar en el programa.

H1: Existe diferencia entre los puntajes de comprensión de lectura antes y después de participar en el programa.

1. Nivel de significación: α = 0.05
2. Realizar el análisis exploratorio de datos y determinación de las medidas descriptivas de las variables antes y después de participar del programa.



1. Generar la variable diferencias ***(di) = Antes - Después***
2. Verificación de Supuestos de Normalidad para las diferencias (di)
   1. Determinar las diferencias ***(di)***
   2. Realizar la prueba de normalidad para las diferencias ***(di)***

Ho: Los datos de las diferencias provienen de poblaciones normales

H1: Los datos de las diferencias no provienen de poblaciones normales

****

H1: Los datos de las diferencias no provienen de poblaciones normales.

1. Dado que lo di, no cumplen con el supuesto de normalidad entonces para comparar los dos grupos se va a utilizar la prueba no paramétrica de Wilcoxon (prueba no paramétrica).
2. Estadística de Prueba No Paramétrica, para dos muestras relacionadas: Prueba de Wilcoxon.

****

****

1. Decisión: Como p<0.05 rechazamos Ho.