

**DISEÑO CENTRADO EN EL HUMANO
Y EXPERIENCIA DE USUARIO**



Entregable 4 Apartado 2:

Evaluación de la usabilidad del Sistema Interactivo de Reserva de las Instalaciones
Deportivas de la UAM

Integrantes:

Qian, Zhije

Ibáñez González, Miguel

Palacio Abrego, Rodolfo

Contraste de hipótesis

Planteamiento de un contraste de hipótesis

Se evaluará si existe una diferencia significativa en el tiempo de realización de la misma tarea entre diferentes grupos de usuarios, sea experto o principiante. El análisis se realizará a través de siguientes medidas:

- Intersujeto: se realiza una comparación con los datos reales y simulados, para medir si el tiempo en completar las tareas entre diferentes grupos, tiene una media parecida o igual.
- Intrasujeto: se realiza una comparación de tiempos con los datos reales para realizar un contraste de si en la Tarea1 (búsqueda y reserva) existe una media o correlación de tiempos con la Tarea 3 (búsqueda por filtro).
- Chi-cuadrado: se intenta medir si los usuarios finales han podido completar las tareas en el número de clic deseado.

El objetivo final de estas medidas es intentar crear unos diseños donde independientemente del perfil de los diferentes usuarios finales, tenga un tiempo parecido en completar las tareas e intentar terminar la tarea con el menor número de clics posibles.

Para poder medir esto, se utilizará intersujetos y Chi-cuadrado para contrastar los resultados producidos, y al tener sólo 5 usuarios finales reales, se simularán 40 datos con la misma media y desviación estándar para formar nuestro primer grupo de datos, y también se intentarán simular otros 45 datos con la misma desviación estándar, pero con una media diferentes para formar nuestro segundo grupo de datos.

Finalmente, para medir que todos los usuarios tengan una curva de aprendizaje parecida, se realizará una medida Intrasujeto para la Tarea 1 y 3, en la Tarea 1 se realiza una primera búsqueda y se intenta reservar una actividad, y en la Tarea 3, se realiza una segunda búsqueda, pero a través de un filtro.

Resultados del análisis

T1-Inter:

En esta primera tarea, se ha producido una media de 23.2 segundos y una desviación de 6.3 segundos con los usuarios finales reales, esto quiere decir, que antes de realizar el contraste de hipótesis, si existiese otro grupo que produjese un intervalo de confianza disjunta al intervalo de confianza muestral, podríamos rechazar de antemano que las medias no son iguales, y el tiempo que tarda en realizar la tarea entre un grupo u otro, también son totalmente desiguales. Pero como tenemos una desviación un poco alta, seguramente existirá un intervalo no disjunto y necesitaríamos realizar una prueba T para poder rechazar la hipótesis, que, en este caso, se ha generado unos datos con una desviación igual pero una media diferente.

En el Excel, se puede ver que simulando con una media de 24 para el grupo 2, fácilmente podemos rechazar la hipótesis nula, esto quiere decir, que existe una diferencia de tiempo entre diferentes grupos y el diseño que se ha planteado no ha conseguido un equilibrio entre distintos perfiles.

	Real + Simulado	Simulado				
Medias	21.20	24.22	Condición: distribuciones normales con varianzas iguales			
			Hipótesis nula: las medias son iguales			
Prueba-t	0.0308815389155672		Función "PRUEBA.T" (significación=0.05; cuarto parámetro=			
Significación	5%		Rechazamos hipótesis nula (HA HABIDO SUERTE)			
Confianza	95%					
Mean_1	23.2					
Std_1	6.30079360080935					
Mean_2	24					
Std_2	6.30079360080935					

T2-Inter:

En esta segunda tarea, se ha producido una media de 26.8 segundo y una desviación alta de 9.5 segundos, una desviación mucho mayor que la Tarea 1. Esto quiere decir que es muy poco probable que haya un intervalo disjunto, por lo que se obliga a utilizar una prueba T para verificar los resultados.

Mirando en el Sheet "T2-Inter_No_rechazo", se puede ver que, aunque hemos decrementado un segundo de media, casi nunca se puede rechazar la hipótesis, ya que al tener una desviación alta, la mayoría de las personas se realizarán las tareas en el mismo rango de tiempos.

	Real + Simulado	Simulado					
Medias	26.31	27.70	Condición: distribuciones normales con varianzas iguales				
			Hipótesis nula: las medias son iguales				
Prueba-t	0.504386964582709		Función "PRUEBA.T" (significación=0,05; cuarto-pa				
Significación	5%		No podemos rechazar hipótesis nula				
Confianza	95%						
Mean_1	26.8						
Std_1	9.49736805646701						
Mean_2	25.8						
Std_2	9.49736805646701						

Tras mover un poquito más la media, con las medidas que se ve en el Sheet “T2-Inter_Rechazo”, ya comenzamos a rechazar la hipótesis, eso quiere decir, que para una media de 28 segundos y con la misma desviación, empieza a notar la diferencia de completar la tarea en diferentes grupos de usuarios.

	Real + Simulado	Simulado					
Medias	26.31	30.60	Condición: distribuciones normales con varianzas iguales				
			Hipótesis nula: las medias son iguales				
Prueba-t	0.0271053373757023		Función "PRUEBA.T" (significación=0,05; cuarto-parámetro=2 p				
Significación	5%		Rechazamos hipótesis nula (HA HABIDO SUERTE)				
Confianza	95%						
Mean_1	26.8						
Std_1	9.49736805646701						
Mean_2	28						
Std_2	9.49736805646701						

T3-Inter:

En esta tercera tarea, sucede algo parecido a lo que ocurre en la tarea 2, la media es de 31.4 segundos y una desviación de 15 segundos. Al final como esperabamos se empieza a rechazar la hipótesis cuando la media del grupo2 sea superior a 34 segundos, con muy pocos rechazos.

	Real + Simulado	Simulado					
Medias	32.48	41.70	Condición: distribuciones normales con varianzas iguales				
			Hipótesis nula: las medias son iguales				
Prueba-t	0.00241245147731811		Función "PRUEBA.T" (significación=0,05; cuarto-parámetro=2 p				
Significación	5%		Rechazamos hipótesis nula (HA HABIDO SUERTE)				
Confianza	95%						
Mean_1	31.4						
Std_1	15.1756383720752						
Mean_2	34						
Std_2	15.1756383720752						

T1-T3-Intra:

En esta medida, se intenta medir si la curva de aprendizaje es igual o parecido para 2 tareas distintas pero parecidas.

En el Sheet “T1-T3-Intra”, utilizando los 5 datos reales, se puede ver que no podemos rechazar la hipótesis nula, eso quiere decir, que posiblemente, ambas tareas tienen una complejidad parecida y los usuarios tienden a realizar ambas tareas con misma escala de tiempos.

Participante	T1 (seg)	T3 (seg)	Diferencia		T1 (seg)	T3 (seg)	Diferencia				
1	23.0	29.0	-6	Medias	23	31	-8	Condición: la diferencia tiene una distribución normal			
2	13.0	26.0	-13					Hipótesis nula: la media de la diferencia es 0			
3	26.0	10.0	16								
4	30.0	45.0	-15	Prueba-t	0.283732973036305			Función "PRUEBA.T" (significación=0.05; cuarto-p			
5	24.0	47.0	-23	Significación	5%			No podemos rechazar hipótesis nula			
6				Confianza	95%						

Luego en el Sheet “T1-T3-Intra2”, se ha utilizado una mezcla de los 5 datos reales con los 40 datos simulados, y esta vez, que se podía prevenir el resultado, se ha rechazado la hipótesis. Esto es bastante normal ya que la mayoría de los datos han sido simulados y existe una desviación alta de tiempos.

Participante	T1 (seg)	T3 (seg)	Diferencia		T1 (seg)	T3 (seg)	Diferencia				
1	23.0	29.0	-6	Medias	21	32	-8	Condición: la diferencia tiene una distribución normal			
2	13.0	26.0	-13					Hipótesis nula: la media de la diferencia es 0			
3	26.0	10.0	16								
4	30.0	45.0	-15	Prueba-t	8.10993861151574E-06			Función "PRUEBA.T" (significación=0.05; cuarto-parámetro=1)			
5	24.0	47.0	-23	Significación	5%			Rechazamos hipótesis nula (HA HABIDO SUERTE)			
6	21.7	14.0		Confianza	95%						

Chi-cuadrado:

En el apartado, se ha realizado diferentes pruebas en el mismo Sheet del Excel.

- Mínimo de clics para Tarea1:

No se ha podido rechazar la hipótesis porque la mayoría de los usuarios finales han conseguido realizar la tarea en mínimo de clics posibles.

Mínimo de click T1	Observado	Esperado									
User 1	9	9									
User 2	9	9									
User 3	12	9		Prueba- chi-cuadrado	0.892505321147118			Medir número mínimo de click			
User 4	10	9		Significación	5%			Hipótesis nula: los usuarios consiguen completar la tarea en el mínimo de clics posibles			
User 5	9	9		Confianza	95%			No podemos rechazar hipótesis nula			

- Mínimo de clics para Tarea 2:

Se ha rechazado la hipótesis nula porque la tarea 2 dispone de diferentes formas de pago, y según la forma elegida, se necesitará menos o más clics para completar la tarea.

Mínimo de click T2	Observado	Esperado
User 1	6	5
User 2	10	5
User 3	10	5
User 4	5	5
User 5	10	5

Prueba chi cuadrado

0.00430388232758925

Significación

5%

Confianza

95%

Hipótesis nula: los usuarios consiguen completar la tarea con el

Rechazamos hipótesis nula (HA HABIDO SUERTE)

Antes de realizar el análisis del apartado anterior, podemos medir el promedio de clics que se necesita para diferentes formas de pago, ya que, al ser un pago, casi siempre se pedirá los mismos datos y será imposible reducir el número de clics. Tras ver el resultado, efectivamente no podemos rechazar la hipótesis ya que la mayoría de las personas, parecen haber completado la tarea con un numero clics parecidos.

Promedio de click T2	Observado	Esperado
User 1	6	7
User 2	10	7
User 3	10	7
User 4	5	7
User 5	10	7
Prueba chi-cuadrado		0.334161717571031
Significación		95%
Confianza		95%

Hipótesis nula: los usuarios consiguen completar la tarea.

No podemos rechazar hipótesis nula

Como la tarea 3 se basa con búsqueda de filtro y cada usuario puede tener su criterio de filtro, se medirá el promedio de clics necesario para encontrar su objetivo. En este caso, no se puede rechazar la hipótesis porque la mayoría de los usuarios finales han conseguido completar la tarea.

[illegible]

Aunque en la medida Intrasujeto ya se ha demostrado que posiblemente no podemos rechazar la hipótesis nula utilizando los 5 datos de usuarios finales, en la medida Chi-cuadrado vuelve a demostrar que no podemos rechazar la hipótesis nula, y posiblemente, la curva de aprendizaje en ambas tareas tiene una escala de tiempo muy parecido.

Observado	T1	T3									
User 1		9	10								
User 2		9	8								
User 3		12	9		Prueba chi-cuadrado	0.175599347955318					
User 4		10	12		Significación	5%		Hipótesis nula: el número de click necesario p:			
User 5		9	15		Confianza	95%		No podemos rechazar hipótesis nula			
Esperado	T1	T3									
User 1		9	9								
User 2		9	9								
User 3		9	9								
User 4		9	9								
User 5		9	9								