

## EJERCICIO PRÁCTICO 6: INFERENCIAS CON FRECUENCIAS – EL CASO DE LAS PROPORCIONES

### CONTEXTO

Ya conocemos bien el contraste de hipótesis sobre la media de una población o las diferencia entre las medias de dos poblaciones. Incluso tenemos herramientas para buscar un nivel de significación ( $\alpha$ ) y un poder estadístico ( $1 - \beta$ ) específicos.

Pero hay muchas preguntas que no se tratan sobre medias, puesto que involucran variables categóricas. Por ejemplo ¿cuánta gente está a favor de la ley de aborto?, ¿qué lista gana la próxima elección del CEII? ¿Cómo podemos hacer frente a este tipo de preguntas?

### OBJETIVOS DE APRENDIZAJE

1. Entender diferentes casos de inferencia estadística con proporciones.
2. Realizar, utilizando el entorno R, inferencias o diseño de estudios con proporciones.

### ÉXITO DE LA ACTIVIDAD

1. El equipo interpreta adecuadamente las preguntas planteadas, identificando el tipo de inferencia o diseño que se consulta.
2. El equipo construye un script en el entorno R para responder adecuadamente cada una de las preguntas planteadas.

### ACTIVIDADES

Los siguientes datos se basan en los resultados de una encuesta publicada por Statista.com (2019) acerca de los adultos estadounidenses y su fanatismo por Star Wars.

Nivel de fanatismo	Mujeres	Hombres
Fanático extremo	158	422
Fanático moderado	291	502
No fanático	431	396

1. Copiar los enunciados de los problemas asignados como comentarios de un script R.
2. Identificar qué se pide determinar en cada pregunta, justificando explícitamente con comentarios en el script.
3. Escribir código R para aplicar la solución propuesta, comentando el procedimiento adecuadamente.
4. Redactar respuestas a las preguntas planteadas (comentarios) en base a los resultados del análisis realizado.

Fuera del horario de clases, cada equipo debe subir el script realizado UVirtual con el nombre "EP06-respuesta-grupo-i", donde i es el número de grupo asignado. Las respuestas deben subirse antes de las 23:30 del lunes 29 de mayo.

## **PREGUNTAS**

### **Grupo 1:**

1. Estudios previos habían determinado más de 30% de las mujeres encuestadas son fanáticas moderadas. ¿Respaldan estos datos tal estimación?
2. Según estos datos, ¿es menor la proporción de fanáticos extremos entre mujeres que entre hombres?
3. Existe la creencia de que hay más fanáticos extremos entre los hombres que entre las mujeres y que dicha diferencia supera el 25%. ¿A cuántas personas (hombres y mujeres) se debería encuestar para obtener un intervalo de confianza del 95% y poder estadístico de 95%, si se intenta mantener aproximadamente la misma proporción de gente estudiada en cada caso?

### **Grupo 2:**

1. Estudios previos habían determinado que menos de 40% de los hombres encuestados son fanáticos moderados. ¿Respaldan estos datos tal estimación?
2. Según estos datos, ¿es igual la proporción de fanáticos (sean moderados o extremos) en hombres y en mujeres?
3. Existe la creencia de que hay más fanáticos extremos entre los hombres que entre las mujeres y que dicha diferencia supera el 25%. ¿A cuántas personas (hombres y mujeres) se debería encuestar para obtener un intervalo de confianza del 97,5% y poder estadístico de 80%, si se intenta mantener aproximadamente la misma proporción de gente estudiada en cada caso?

### **Grupo 3:**

1. Estudios previos habían determinado que 45% de las mujeres encuestadas no son fanáticas. ¿Respaldan estos datos tal estimación?
2. Según estos datos, ¿es mayor la proporción de fanáticos moderados entre hombres que entre mujeres?
3. Existe la creencia de que hay más fanáticos extremos entre los hombres que entre las mujeres y que dicha diferencia supera el 25%. ¿A cuántas personas (hombres y mujeres) se debería encuestar para obtener un intervalo de confianza del 99% y poder estadístico de 75%, si se intenta mantener aproximadamente la misma proporción de gente estudiada en cada caso?

### **Grupo 4:**

1. Estudios previos habían determinado que más de 32% de las mujeres encuestadas son fanáticas moderadas. ¿Respaldan estos datos tal estimación?
2. Según estos datos, ¿es menor la proporción de fanáticos extremos entre mujeres que entre hombres?
3. Existe la creencia de que hay más fanáticos extremos entre los hombres que entre las mujeres y que dicha diferencia supera el 30%. ¿A cuántas personas (hombres y mujeres) se debería encuestar para obtener un intervalo de confianza del 95% y poder estadístico de 95%, si se intenta mantener aproximadamente la misma proporción de gente estudiada en cada caso?

### **Grupo 5:**

1. Estudios previos habían determinado que menos de 42% de los hombres encuestados son fanáticos moderados. ¿Respaldan estos datos tal estimación?
2. Según estos datos, ¿es igual la proporción de fanáticos extremos en hombres y en mujeres?
3. Existe la creencia de que hay más fanáticos extremos entre los hombres que entre las mujeres y que dicha diferencia supera el 30%. ¿A cuántas personas (hombres y mujeres) se debería encuestar para obtener un

intervalo de confianza del 97,5% y poder estadístico de 80%, si se intenta mantener aproximadamente la misma proporción de gente estudiada en cada caso?

**Grupo 6:**

1. Estudios previos habían determinado que 47% de las mujeres encuestadas no son fanáticas. ¿Respaldan estos datos tal estimación?
2. Según estos datos, ¿es mayor la proporción de fanáticos extremos entre hombres que entre mujeres?
3. Existe la creencia de que hay más fanáticos extremos entre los hombres que entre las mujeres y que dicha diferencia supera el 30%. ¿A cuántas personas (hombres y mujeres) se debería encuestar para obtener un intervalo de confianza del 99% y poder estadístico de 75%, si se intenta mantener aproximadamente la misma proporción de gente estudiada en cada caso?

**Grupo 7:**

1. Estudios previos habían determinado que más de 27% de los hombres encuestados no son fanáticos. ¿Respaldan estos datos tal estimación?
2. Según estos datos, ¿es menor la proporción de no fanáticos entre hombres que entre mujeres?
3. Existe la creencia de que hay más fanáticos extremos entre los hombres que entre las mujeres y que dicha diferencia supera el 35%. ¿A cuántas personas (hombres y mujeres) se debería encuestar para obtener un intervalo de confianza del 95% y poder estadístico de 95%, si se intenta mantener aproximadamente la misma proporción de gente estudiada en cada caso?

**Grupo 8:**

1. Estudios previos habían determinado que menos de 35% de las mujeres encuestadas son fanáticas moderadas. ¿Respaldan estos datos tal estimación?
2. Según estos datos, ¿es igual la proporción de no fanáticos en hombres y en mujeres?
3. Existe la creencia de que hay más fanáticos extremos entre los hombres que entre las mujeres y que dicha diferencia supera el 35%. ¿A cuántas personas (hombres y mujeres) se debería encuestar para obtener un intervalo de confianza del 97,5% y poder estadístico de 80%, si se intenta mantener aproximadamente la misma proporción de gente estudiada en cada caso?

**Grupo 9:**

1. Estudios previos habían determinado que 35% de los hombres encuestados son fanáticos extremos. ¿Respaldan estos datos tal estimación?
2. Según estos datos, ¿es mayor la proporción de no fanáticos entre hombres que entre mujeres?
3. Existe la creencia de que hay más fanáticos extremos entre los hombres que entre las mujeres y que dicha diferencia supera el 35%. ¿A cuántas personas (hombres y mujeres) se debería encuestar para obtener un intervalo de confianza del 99% y poder estadístico de 75%, si se intenta mantener aproximadamente la misma proporción de gente estudiada en cada caso?

**Grupo 10:**

1. Estudios previos habían determinado que más de 29% de los hombres encuestados no son fanáticos. ¿Respaldan estos datos tal estimación?
2. Según estos datos, ¿es menor la proporción de fanáticos moderados en hombres y en mujeres?

3. Existe la creencia de que hay más fanáticos extremos entre los hombres que entre las mujeres y que dicha diferencia supera el 15%. ¿A cuántas personas (hombres y mujeres) se debería encuestar para obtener un intervalo de confianza del 95% y poder estadístico de 95%, si se intenta mantener aproximadamente la misma proporción de gente estudiada en cada caso?

## CRITERIOS DE EVALUACIÓN

- Obtienen los datos adecuados para responder cada una de las preguntas del ejercicio, identificando las frecuencias involucradas y determinando correctamente los totales y proporciones involucradas.
- Formulan con claridad y explícitamente hipótesis nula y alternativa adecuadas para responder la pregunta planteada sobre una proporción, tanto en lenguaje natural como en lenguaje matemático, identificando clara y explícitamente la prueba que puede contrastarlas.
- Escriben código R que verifica, o donde se argumenta convincentemente, el cumplimiento de las condiciones de la prueba, y que realiza de forma completa y correcta la prueba seleccionada con los datos correctos.
- Entregan una conclusión correcta y completa a la pregunta planteada, basándose en el resultado de la prueba realizada y el contexto del problema.
- Formulan con claridad y explícitamente hipótesis nula y alternativa adecuadas para responder la pregunta planteada sobre la diferencia entre dos proporciones, tanto en lenguaje natural como en lenguaje matemático, identificando clara y explícitamente la prueba que puede contrastarlas.
- Escriben código R que verifica, o donde se argumenta convincentemente, el cumplimiento de las condiciones de la prueba, y que realiza de forma completa y correcta la prueba seleccionada con los datos correctos.
- Entregan una conclusión correcta y completa a la pregunta planteada, basándose en el resultado de la prueba realizada y el contexto del problema.
- Reconocen que se pregunta por el tamaño de dos muestras, identificando claramente el contexto de la hipótesis nula que se desea contrastar, la prueba que necesita realizarse y los elementos conocidos (alfa, beta, tamaño del efecto, etc.).
- Obtienen correctamente los tamaños de las muestras para una prueba de diferencia de proporciones, considerando los elementos conocidos (alfa, beta, tamaño del efecto, etc.) y manteniendo aproximadamente la proporción de personas estudiadas en cada caso.
- Escriben código R -ordenado, bien indentado, sin sentencias espurias y bien comentado- que intenta dar respuesta a todas las preguntas planteadas.
- Escriben con buena ortografía y redacción ( $\leq 3$  errores), usando vocabulario propio de la disciplina y el contexto del problema.