



Actividad: 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 (α) 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 CEI? ¿Cómo podemos hacer esto?

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 puede responder adecuadamente cada una de las preguntas planteadas usando el entorno R

Actividades.

1. El profesor introduce un poco de teoría del estadístico (20 min.)
2. El profesor sortea las preguntas (5 min.)
3. El equipo lee el enunciado, lo discute, calcula (en R) las condiciones para lograr una respuesta e interpreta este resultado para responder cada una de las preguntas entregadas (45 min.)
4. El equipo comparte con el resto del curso su trabajo (20 min.)

Enunciado y datos para las preguntas

Los siguientes datos vienen de un estudio hecho por Rothman & Keller (1972) (Journal of chronic diseases, 25(12), 711-716) sobre la incidencia de la cantidad de alcohol y de tabaco que se consume en el riesgo de padecer cáncer oral.

Alcohol (ml por día)	Casos de cáncer oral	Controles
0	43	108
1-9	89	141
10-44	109	91
45+	242	107

Tabaco (cigarrillos por día)	Casos de cáncer oral	Controles
0	26	85
1-19	66	97
20-39	248	197
40+	143	68



Preguntas A

1. Suponiendo que la diferencia en la proporción de personas que desarrollan la enfermedad entre quienes beben de 10 a 44 ml de alcohol por día y aquellos que beben 45 o más ml al día es de 0.15. ¿Cuánta gente deberíamos entrevistar para obtener un intervalo de confianza del 95% y poder estadístico de 90%?
2. Estudios previos habían determinado que la incidencia de cáncer oral en la población general que bebe regularmente entre 10 y 44 ml de alcohol era de 50%. ¿Respaldan estos datos tal estimación?
3. Según estos datos, ¿da lo mismo beber de 10 a 44 ml de alcohol diariamente que hacerlo con 45 o más ml?

Preguntas B

1. Suponiendo que la diferencia en la proporción de personas que desarrollan la enfermedad entre quienes beben de 1 a 9 ml de alcohol por día y aquellos que beben de 10 a 44 ml al día es de 0.10. ¿Cuánta gente deberíamos entrevistar para obtener un intervalo de confianza del 95% y poder estadístico de 80%?
2. Estudios previos habían determinado que la incidencia de cáncer oral en la población que bebe regularmente entre 1 y 9 ml de alcohol era de 25%. ¿Respaldan estos datos tal estimación?
3. Según estos datos, ¿da lo mismo beber entre 1 y 9 ml de alcohol diariamente que hacerlo de 10 a 44 ml?

Preguntas C

1. Suponiendo que la diferencia en la proporción de personas que desarrollan la enfermedad entre quienes no fuman y aquellos que fuman de 1 a 19 cigarrillos al día es de 0.2. ¿Cuánta gente deberíamos entrevistar para obtener un intervalo de confianza del 95% y poder estadístico de 80%?
2. Estudios previos habían determinado que la incidencia de cáncer oral en la población general que no fuma era de 15%. ¿Respaldan estos datos tal estimación?
3. Según estos datos, ¿da lo mismo no fumar que hacerlo diariamente consumiendo entre 1 y 19 cigarrillos?

Preguntas D

1. Suponiendo que la diferencia en la proporción de personas que desarrollan la enfermedad entre quienes fuman de 20 a 39 cigarrillos al día y aquellos que fuman 40 o más unidades por día es de 0.18. ¿Cuánta gente deberíamos entrevistar para obtener un intervalo de confianza del 95% y poder estadístico de 80%?
2. Estudios previos habían determinado que la incidencia de cáncer oral en la población general que fuma entre 20 y 39 cigarrillos al día era de 5%. ¿Respaldan estos datos tal estimación?
3. Según estos datos, ¿da lo mismo fumar diariamente entre 1 y 2 paquetes de cigarrillos que hacerlo más de dos paquetes?