



DEPARTAMENTO DE INGENIERÍA CIVIL INDUSTRIAL
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
UNIVERSIDAD DE CHILE
IN3242-1 ESTADÍSTICA

LABORATORIO N°1

Integrantes: Adolfo Rojas
Profesor: Daniel Schwartz P.
Auxiliar: Antonia Aceituno
Camila Galarce
Ayudantes: Bastián Medina
Guillermo Escobar Fuentes

Fecha de entrega: 27 de agosto de 2024
Santiago de Chile

- El valor dado de solo considerar una única variable explicativa (es decir un solo regresor, ya sea el asociado al tiempo de visita o cualquier otra pues no se especifica) indica la proporción de simulaciones donde dicho regresor es significativo al 0.05 para una única regresión y el número total de simulaciones
- Ahora se indica la proporción entre las simulaciones donde en **al menos una** de las 5 regresiones se de que el regresor (estimador de $\hat{\beta}_1$) de la variable explicativa sea significativa (distinta de 0) y el número total de simulaciones
- Con Chat gpt creamos un vector que guarde las proporciones para cada loop (cada iteración del loop va agregando regresiones/var. dependientes) y el gráfico que se ve en la fig 1, la conclusión que podemos sacar es que a medida de que se empiecen a medir más cosas (hacer varias regresiones) se de un p-value que indique significancia (al 0.05) cuando en un comienzo (con una única regresión) probablemente no exista y de que el regresor no sea = 0 cuando en verdad no se puede aceptar la hipótesis alternativa

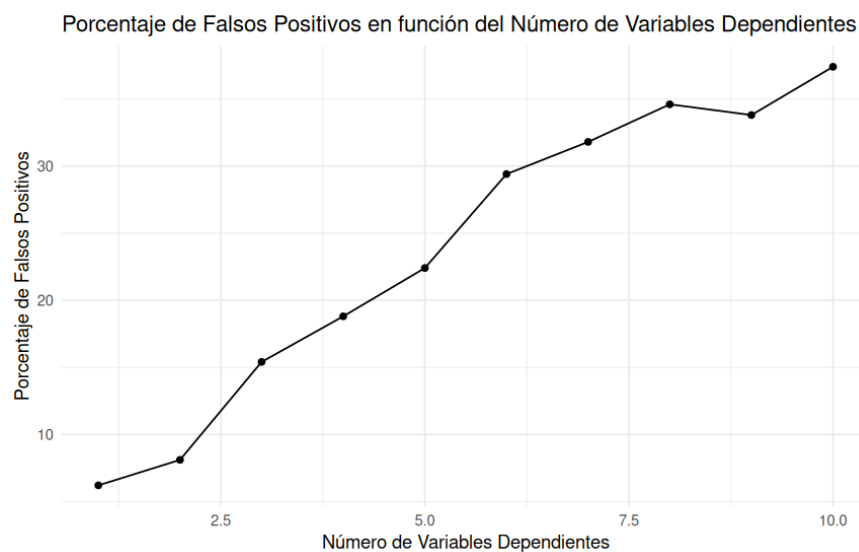


Figura 1: Porcentaje de falsos positivos según la cantidad de regresiones

- En el ítem anterior se contesta en gran medida lo pedido en este ítem, nosotros calculamos el p-valor del estimador del regresor de la única variable explicativa de cada modelo y vemos como a medida de que aumentamos la cantidad de regresiones donde se considera esta variable se vuelve más complicado rechazar la hipótesis inicial (de que el regresor sea ~ 0 y se pueda omitir la variable pues no es significativa)
- Bonus: A medida de que aumenten la cantidad de participantes y simulaciones uno esperaría que aumente a su vez la precisión con la cual se estima el regresor (bajando su error)