Pontificia Universidad Católica de Chile
Instituto de Astrofísica — Facultad de Física
Estadística para Astrónomos - ASP5408
Primer Semestre 2016

Tarea 1

Profesor: Andrés Jordán (ajordan@astro.puc.cl)
Ayudante: Néstor Espinoza (nespino@astro.puc.cl)

Fecha de entrega: Viernes 1ro de Abril, 11:30 hrs. (buzón del ayudante y/o en webcursos).

La presente tarea tiene por objetivo el ejercitar sus habilidades matemáticas aplicadas al área de probabilidades y estadística, asi como también en el lenguaje de programació Python.

A considerar:

- Cualquier código que haya sido ocupado para desarrollar la tarea debe adjuntarse en la resolución de la misma a través de un link a su GitHub. Si no tiene cuenta en GitHub, cree una.
- Desarrolle lo más que pueda sus ideas.
- Esta bien discutir las tareas en grupo, pero recuerde que estas son personales. Usted debe generar una discusión propia de sus resultados.

Problema 1 Preguntas cortas

a) ¿Cuál es la diferencia entre eventos disjuntos y eventos independientes? De un ejemplo astronómico de un evento disjunto y un evento independiente.

b) Suponga que dos eventos, A y B, son independientes. Pruebe que A^c y B^c son independientes.

. . . .

Problema 2 Estudiando encuestas

En clases vimos el famoso problema de Monty Hall, y discutimos como la mayoría de la gente elige quedarse con su respuesta inicial en vez de cambiarse (que, como vimos, es la mejor estrategia) una vez que Monty Hall revela una puerta sin el premio. Para corroborar esta última hipótesis, el ayudante del curso generó una encuesta en Twitter[1], la que fue respondida por 33 personas: 18 de ellas decidieron quedarse con la puerta original, mientras que 15 decidieron cambiarse de puerta. En este problema intentaremos responder si, efectivamente, podemos concluir probabilísticamente que la mayoría de la gente escoge la peor estrategia como habíamos comentado en clase.

a) Observe que en este caso hay dos respuestas posibles. Si suponemos la gente escoge la primera con probabilidad r, y la segunda con probabilidad $1-r$, y nosotros estamos interesados en la distribución de X, el número de veces que la gente escoge la primera respuesta, ¿qué distribución tendría entonces X?, ¿qué parámetros observados tenemos de dicha distribución?

b) ¿Qué tan raro es observar los datos que observamos si asumimos que $r = 0.5$? Calcule primero la probabilidad de observar $X = 18$, como observamos con nuestros datos, i.e., $\mathbb{P}(X = 18|r = 0.5)$. Luego, escriba un código que simule el proceso de votación 1000 veces y calcule cuantas veces observa $X = 18$; ¿son consistentes sus resultados?

c) Note que lo que buscamos es comparar dos hipótesis: $r > 0.5$ (i.e., la mayoría de la gente tiende a escoger la primera respuesta) v/s $r < 0.5$ (la mayoría de la gente escoge la segunda). Para compararlas, necesitamos la distribución de p dada nuestra observación $X = x$, donde $x = 18$ en este caso, además de el(los) parámetro(s) que ud. describió en a). Note que esto se puede obtener gracias al Teorema de Bayes. Escriba usando este teorema la PDF $p(r|X)$, la distribución de r dados nuestros datos observados. Describa cada término que aparece usando dicho teorema. En particular, observará que entre dichos términos aparecerá $p(r)$; asuma este termino es igual a 1 y explique qué significa esto (dibuje dicha distribución, intégrela en el rango posible para r, etc.). (i) Grafique la distribución obtenida (note que necesitará un integrador numérico para calcular el denominador); (ii) ¿dónde está el máximo de esta distribución? y (iii) grafique la CDF de su distribución (también necesitará un integrador numérico para esto) (iv) verifique que su distribución integra a 1 en todo el rango posible de r.

d) Usando un integrador numérico y su resultado obtenido en b), calcule $\mathbb{P}(r > 0.5|X)$ y $\mathbb{P}(r < 0.5|X)$ y compárelas; ¿qué puede concluir?

....

[1] `https://twitter.com/nespinozap/status/709356553028169730`