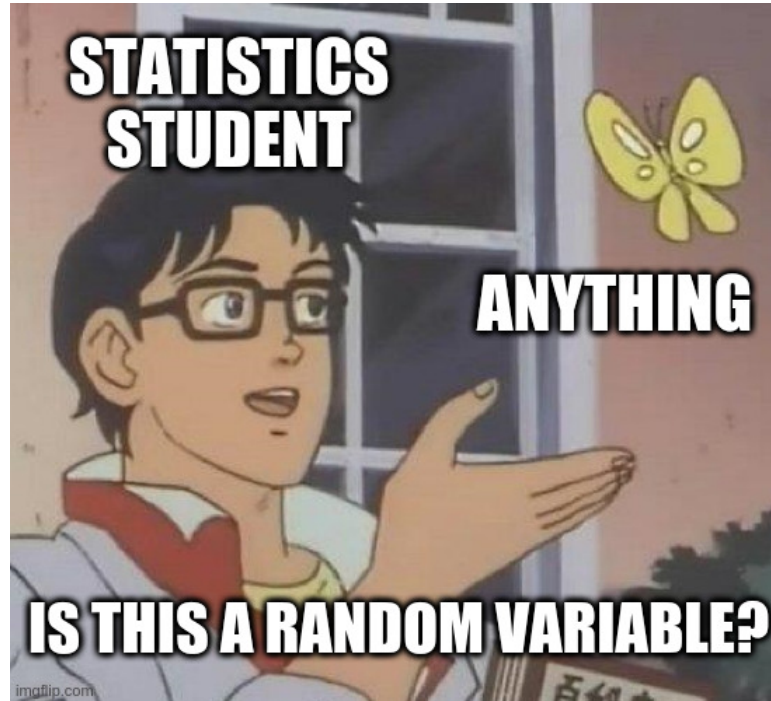


# INF-280: Estadística Computacional



Probabilidad y Variables Aleatorias

# Temario

- Repaso
  - Fundamentales de probabilidad ❤️
  - Resultados importantes
  - Variables aleatorias
  - Distribuciones de probabilidad
- Ejercicios

# Repaso

## Fundamentales de probabilidad

- Experimento aleatorio: Procedimiento bien definido que produce un resultado observable sujeto a incertidumbre.
- Espacio muestral  $\Omega$  : Conjunto de todos los resultados posibles de un experimento aleatorio.
- Evento: Subconjunto de  $\Omega$ .
- Conjunto de Eventos (espacio de eventos)  $\mathcal{F}$ : Conjunto de subconjuntos de  $\Omega$ .

Es el conjunto potencia de  $\Omega$ , su cardinalidad es  $2^{|\Omega|}$

- Medida de probabilidad: Es una función  $P : \mathcal{F} \rightarrow [0, 1] \subset \mathbb{R}$  que satisface los siguientes axiomas (kolgomorov):
  - i.  $P(A) \geq 0, \forall A \in \mathcal{F}$
  - ii.  $P(\Omega) = 1$
  - iii. Si  $A_1, A_2, \dots$  son eventos disjuntos entonces:  $P(\bigcup_i A_i) = \sum_i P(A_i)$

## Resultados importantes

Considerando los eventos  $A$  y  $B$ :

1. Probabilidad condicional: Sea  $P(B) > 0$ , la probabilidad de que ocurra  $A$  dado  $B$  viene dada por:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

2. Ley de probabilidad total: Sea  $A_1, \dots, A_k$  una partición de  $\Omega$ . Entonces, para cualquier evento  $B$  tenemos:

$$P(B) = \sum_{i=1}^k P(B|A_i) \cdot P(A_i)$$

3. Teorema de Bayes: ❤️

$$P(A) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

## Variables aleatorias

- Variable aleatoria: Función inyectiva que asigna a cada elemento  $\omega$  del espacio muestral  $\Omega$  un número  $X(\omega)$ :

$$X : \Omega \rightarrow E$$

Codificación del espacio muestral.

- Función de probabilidad: Asigna a cada estado de una variable aleatoria su respectivo grado de certeza. Puede ser una PDF o una PMF.
- Función de Distribución Acumulada: La función de distribución acumulada  $F_X : \mathbb{R} \rightarrow [0, 1]$  viene dada por:

$$F : P(X \leq a)$$





## Distribuciones de probabilidad

1.  $X \sim \text{Bernoulli}(p)$
2.  $X \sim \text{Geometrica}(p)$
3.  $X \sim \text{Binomial}(n, p)$
4.  $X \sim \text{Poisson}(\lambda)$
5.  $X \sim \text{Uniforme}(a, b)$

# Ejercicios

1. Suponga que tiene dos fuentes llenas de galletas. La fuente 1 contiene 30 galletas de vainilla y 10 de chocolate. La fuente 2 contiene 20 de cada tipo. Ahora, suponga que escoge una de las fuentes al azar y, sin mirar, elige una galleta. La galleta es de vainilla.
  - [a] Defina las variables aleatorias de interés.
  - [b] ¿Cuál es la probabilidad de que haya venido de la fuente 1?
  - [c] ¿Qué ocurre con la probabilidad anterior cuando la probabilidad de elegir una galleta de la primera fuente tiende a 0?

2. Considere el siguiente algoritmo aleatorio que busca adivinar la contraseña *password* de  $n$  dígitos:

```
def guess_pass(password):  
    password_len = len(password)  
    # Generate Bernoulli password guess  
    guess = generate_guess(password_len)  
    # Repeat until guess password  
    while(check_pass(guess, password) == False):  
        guess = generate_guess(n)  
    return guess
```

- [a] ¿Qué distribución sigue el **número de iteraciones** de `guess_pass` ?
- [b] ¿Cuál es el número esperado de iteraciones para adivinar una contraseña de 6 dígitos?

Considere ahora el siguiente algoritmo que realiza el `check_pass` luego de  $n$  intentos:

```
def guess_pass(n, password):  
    password_len = len(password)  
    # Perform n attempts  
    guesses = [generate_guess(password_len) for guess in range(n)]  
    # Check every guess  
    for guess in guesses:  
        if(check_pass(guess, password)):  
            return guess  
    return None
```

- [c] ¿Cuál es la distribución que sigue el número de **intentos exitosos**?
- [d] ¿Cuál es el número esperado de intentos exitosos considerando una contraseña de 6 dígitos y 1000000 intentos?

3. Hay dos contadores en una oficina particular que preparan declaraciones de impuestos. Para un tipo particular de formulario, el número de errores cometidos por el primer contador tiene una distribución de Poisson con media  $\lambda_1$  y el número de errores cometidos por el segundo contador tiene una distribución de Poisson con media  $\lambda_2$ . Cada contador prepara el mismo número de formularios de este tipo y la función de masa de probabilidad del número de errores  $X$  en un formulario seleccionado al azar es:

$$p(x) = 0.5 \frac{e^{-\lambda_1} \lambda_1^x}{x!} + 0.5 \frac{e^{-\lambda_2} \lambda_2^x}{x!} \quad x = 0, 1, 2, \dots$$

- [a] Compruebe que  $p(x)$  es de hecho una función de masa de probabilidad legítima.
- [b] ¿Cuál es el número esperado de errores en el formulario seleccionado?
- [c] ¿Cuál es la varianza del número de errores en el formulario seleccionado?

4. El artículo “Modeling Sediment and Water Column Interactions for Hydrophobic Pollutants” (Water Research, 1984:1169-1174) sugiere la distribución uniforme en el intervalo (7.5, 20) como modelo de profundidad (cm) de la capa de bioturbación en sedimento en una región.
- [a] ¿Cuáles son la media y la varianza de la profundidad?
  - [b] ¿Cuál es la función de distribución acumulativa de la profundidad?
  - [c] ¿Cuál es la probabilidad de que la profundidad observada sea cuando mucho de 10? ¿Entre 10 y 15?
  - [d] ¿Cuál es la probabilidad de que la profundidad observada esté dentro de una desviación estándar del valor medio? ¿Dentro de dos desviaciones estándar?