**Instituto Tecnológico de Costa Rica**

**Escuela de Ingeniería en Computación**

**Tarea #1**

**“Curso de Estadística para el programa de Ciencia de Datos”**

**Andrés González Ugalde**

**Sede San Carlos**

**2019**

**I PARTE**

1. Para cada uno de dos relojes se compara durante varios días, la hora que marcan con respecto a la hora exacta. El primer reloj adelanta un promedio de 3,2 segundos por día, con una desviación estándar de 0,8 segundos. El segundo reloj atrasa un promedio de 5,7 segundos, con una desviación estándar de 0,2 segundos. ¿Cuál de los dos relojes es más confiable? Explique

Calculamos el porcentaje de variación de ambos relojes, de la media con respecto a la desviación estándar, para encontrar cuál de los dos relojes tiene mayor dispersión en sus datos.

**Reloj 1:**

**Reloj 2:**

**R/** El **reloj 2** tiene un porcentaje de dispersión menor a la del **reloj 1**, esto lo que nos indica es que es más común que el **reloj 2**, se atrase una cantidad de segundos cerca de la media comparado al **reloj 1**, que tiende a atrasarse una cantidad de segundos en un rango más variable en relación a la media.

1. Una costumbre reciente en los cines es mostrar anuncios comerciales y cortos en la pantalla antes de comenzar la película. El tiempo de espera para una película es la diferencia entre el tiempo de comienzo anunciado y el tiempo real en que comienza la película. En un grupo de 21 estudiantes investigó el tiempo de espera de dos cines de las cadenas de cine mexicanas: Cinemex y Cinépolis. Cada estudiante asistió a dos funciones, una película en cada cine y registró los tiempos de espera en minutos. Los datos arrojaron los siguientes resultados:

|  |  |  |
| --- | --- | --- |
| **Estadístico** | **Cinemex** | **Cinépolis** |
| Media | 10 | 10 |
| Desviación Estándar | 2,8062 | 1,0368 |

1. Un estudiante en clase argumenta que en realidad no hay diferencia en los tiempos de espera de esos cines, ya que la media es la misma en ambos casos. ¿Es correcto ese dato? ¿Estás de acuerdo con ese estudiante? ¿Por qué? Haga un análisis y una interpretación con las desviaciones estándar

Calculamos el porcentaje de variación de ambos cines, de la media con respecto a la desviación estándar, para encontrar cuál de los dos cines tiene mayor dispersión en sus datos.

**Cinemex:**

**Cinépolis:**

**R/** No es correcto deducir que el tiempo de espera va a ser similar en ambos cines ya que tienen la misma media, dado que la desviación estándar es diferente y esto afecta que tan disperso están los datos de esa media, en este caso Cinépolis tiene un porcentaje de dispersión menor, por lo tanto, es más probable que el tiempo de espera sea cercano a la media.

1. Suponga que eres una persona muy puntual y que no tienes mucha paciencia para estar viendo comerciales, ¿Cuál de esas cadenas de cine elegirías para ver una película? ¿Por qué?

**R/** Si yo soy una persona muy puntual y no tengo la paciencia de esperar viendo comerciales, escogería ir a Cinépolis ya que la dispersión es menor en relación a la media, y es más probable que el tiempo de espera este cerca de la media y me sería más fácil llegar en un tiempo que los comerciales estén por terminar. A diferencia de Cinemex donde hay más incertidumbre sobre el tiempo de espera ya que la dispersión es mayor en relación a la media.

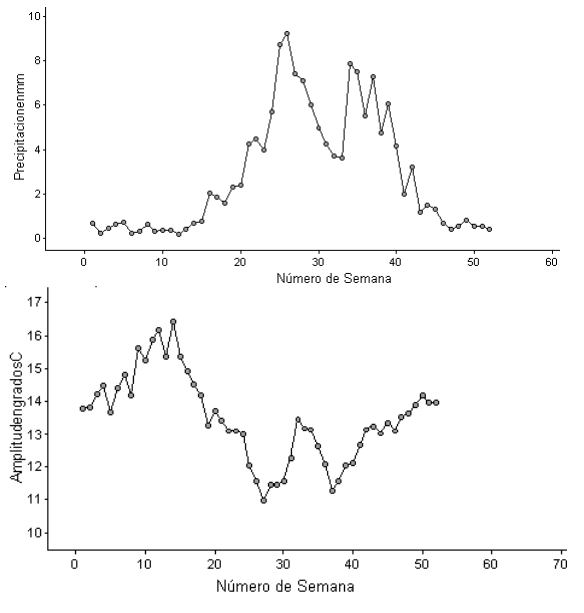
1. Suponga que usted está realizando un tour en una selva y llega a un punto en el que debe cruzar al otro lado de la orilla de un río. Revisando el mapa se identifica una nota que indica que el ese río tiene una profundidad promedio de 1m. Suponiendo que sus dotes de nadador no son muy destacables, ¿se animaría a cruzar el río caminando?, o, por el contrario, antes de hacerlo o de pensar en otra opción, ¿Qué otro tipo de información considera que necesitaría saber? Explique

**R/** Si tuviera el dato de la desviación estándar podría tomar una mejor decisión, si la desviación se aleja más de 1m de la media, estaría en problemas porque podría encontrar que el rio tiene puntos con una profundidad mayor a mi altura, si la desviación es de 0.1 a 0.4 ningún punto del rio superaría mi estatura y podría pasar caminando.

**II PARTE**

Seguidamente se le presentan dos gráficos correspondientes a la variación de temperatura y la precipitación promedio, por semana, de la comunidad de El Mante en el estado de Tamaulipas, México. Estos promedios están basados en datos recolectados durante 38 años en esta región. Usted debe determinar cuál es la época más apropiada para que se realice la zafra (Corta la de la caña de azúcar) tomando como referencia lo observado en los gráficos. No olvide argumentar adecuadamente su respuesta. Para resolver este ejercicio debe contemplar las siguientes restricciones propuestas el investigador Ángel Martínez Garza en 1978:

1. La lluvia y temperatura de la zona cultivada con caña, influye en la concentración de azúcar de la caña cosechada.
2. Cuando hay mucha variación de temperatura un mes antes de la cosecha, se produce un efecto positivo en el rendimiento del azúcar.
3. La lluvia registrada en un periodo de dos meses antes de la cosecha reduce el rendimiento del azúcar.
4. En el caso de la región de El Mante, la duración de la zafra es aproximadamente de 14 semanas



**R/** De la semana 15 a la 20 parece ser un buen rango para comenzar la zafra ya que han pasado 4 semanas de cambio de temperatura y muy pocas semanas con lluvias que afecten la siembra. Las primeras semanas no son recomendables porque no hay cambios de temperaturas significativos que afecten positivamente la siembra, las semanas de por medio tienen muchas lluvias que pueden afectar la siembra negativamente, y las semanas de final de año no tienen cambios de temperatura significativos.