

El problema de las audiencias

Medición del Tiempo de estadía promedio

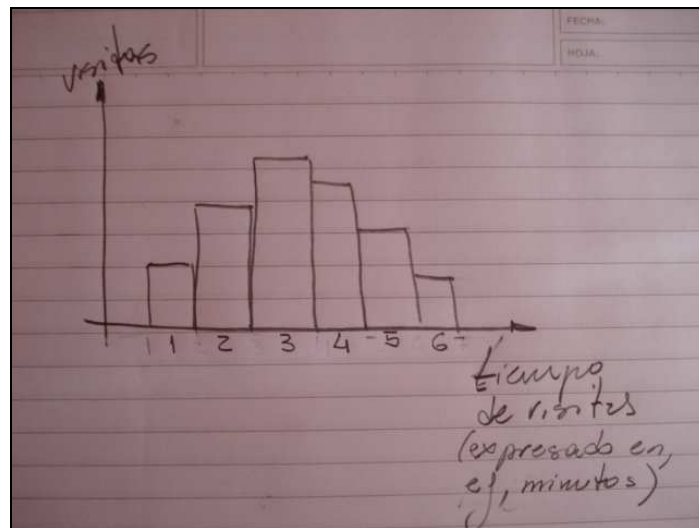
La pregunta es ¿cómo formular las diferencias entre sitios web's o programas de televisión en cuanto a visitas y tiempo de estadía de las mismas?

Una primera aproximación simple viene dada por la siguiente formulación matemática:

$$TPE = \frac{1}{n} \sum_{i=1}^n TIU$$

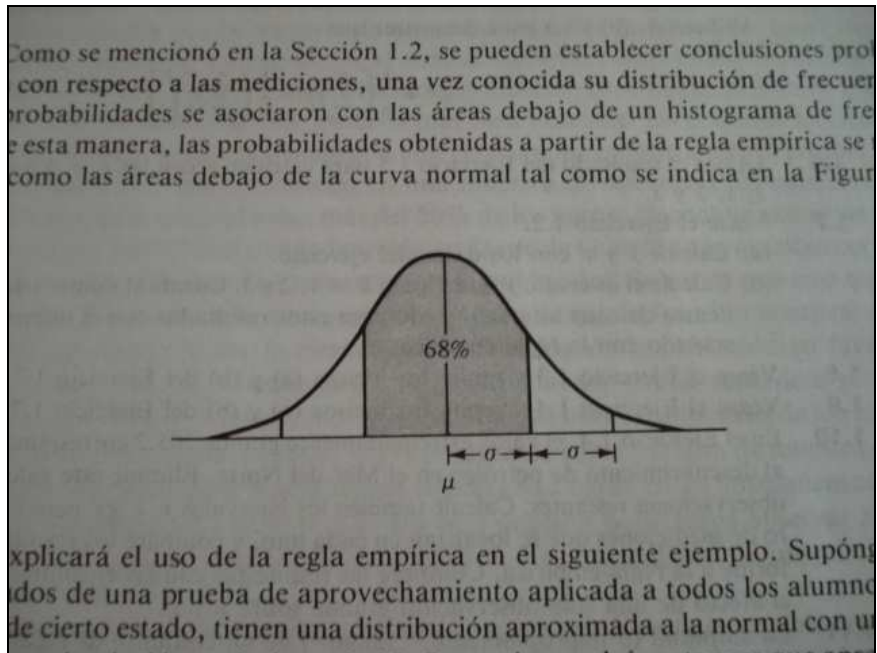
Que quiere decir que el Tiempo Promedio de Estadía (*TPE*) del usuario típico es igual a la sumatoria de los Tiempos Invertidos por cada uno de los Usuarios (*TIU*) dividido el total de visitantes (*n*). **Es decir, lo que habría que hacer es sumar cada uno de los tiempos invertidos por cada uno de los usuarios y dividirlo por el total de usuarios.** Así obtendríamos el tiempo promedio de estadía del usuario típico.

Sin embargo este tratamiento podría ser desarrollado, aunque más no sea una primera vez o una vez cada tanto, junto a este otro. Se debería trazar el siguiente gráfico:



En donde en el eje de las abscisas se trazara el tiempo de estadía de los usuarios y en el de las ordenadas la cantidad de visitas por cada uno de los rangos de tiempo de visitas. ¿Porqué sería útil realizar este paso? Porque nos daría una mejor información de lo que está sucediendo con las visitas y porque podríamos hacer más fuertes aseveraciones. Veamos cómo.

Por un lado este análisis serviría porque podríamos hacer aseveraciones interesantes como la siguiente:

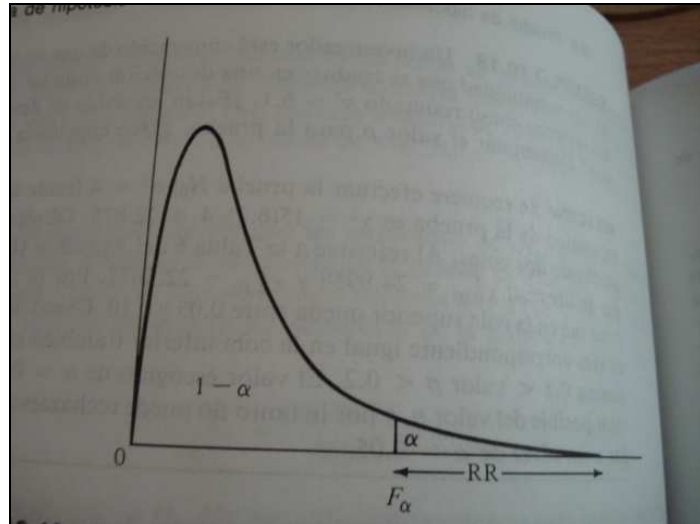


En este caso imaginario podríamos decir: “El sitio Clarín tiene 340.000 visitas por día, que representan el 68% del total de visitas que recibe el sitio por día, que pasan entre 3 y 5 minutos en el sitio cada vez que entran. En promedio unos 4 minutos por usuario típico”. El valor de ver las distribuciones es justamente el de poder aseverar más fuertemente la varianza de la muestra. Es decir, la dispersión de los valores obtenidos. Hay sitios que tendrán más variabilidad de tiempo de estadía (e.g. entre 2 y 10 minutos) y otros menos (entre 3 y 5 minutos). Entonces se podrían hacer comparaciones tales como:

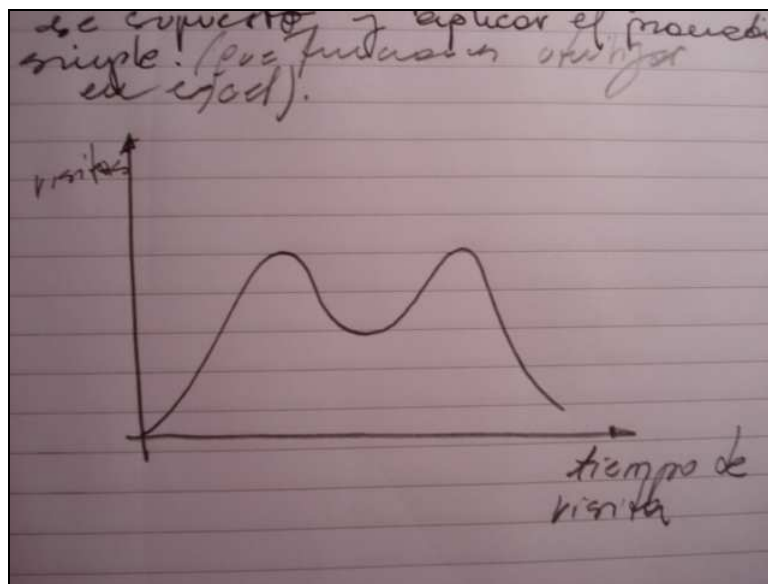
Sitio Clarín: El sitio Clarín tiene 340.000 visitas por día, que representan el 68% del total de visitas que recibe el sitio por día, que pasan entre 3 y 5 minutos en el sitio cada vez que entran. En promedio unos 4 minutos por usuario típico en el sitio.

Sitio Infobae: El sitio Infobae tiene 140.000 visitas por día, que representan el 68% del total de visitas que recibe el sitio por día, que pasan entre 1 y 3 minutos en el sitio cada vez que entran. En promedio unos 2 minutos por usuario típico en el sitio.

Por otro lado esto serviría para los casos, y esto es importante, en que la muestra no se comportase como se señaló arriba. Es decir, cuando la muestra no se pudiese aproximar a una distribución normal. Por ejemplo esto podría suceder en casos como los siguientes:



En donde la distribución estuviera sesgada.



O en los casos en donde la distribución tuviera dos “picos”. En estos casos el tratamiento debiera ser diferente. Se deberían usar, por citar algunos, otro tipo de promedios propios de esas distribuciones, la mediana o el valor más repetido. De lo contrario el promedio simple nos daría información de poco o ningún valor.

Si, como es de suponer, nos encontrásemos con que las distribuciones tendieran a una normal en casi todos los sitios analizados entonces se podría utilizar el promedio expuesto al principio sin tener problema alguno. Entonces de esa manera se podría decir, el sitio Clarín tiene 500.000 visitas que en promedio pasan 4 minutos en la página.

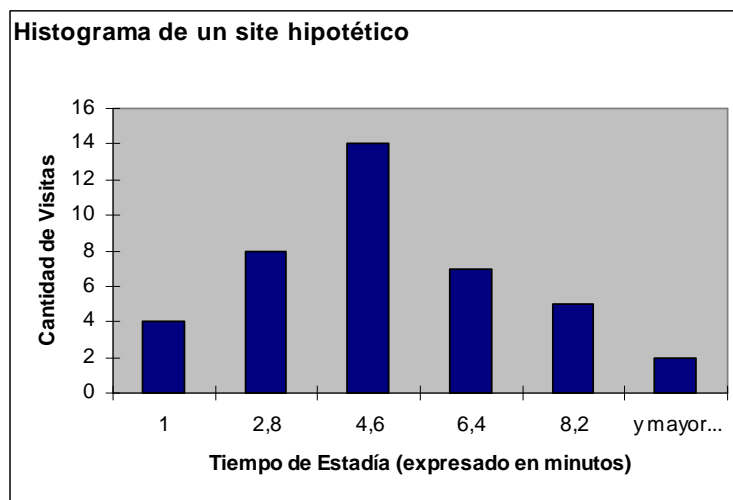
Obsérvese la diferencia de información visualizada entre los dos planteos. El más elaborado expuesto arriba y el más simple expuesto a continuación.

De la forma simple esto quedaría:

Sitio Clarín: El sitio Clarín tiene 500.000 visitas por día que en promedio pasan unos 4 minutos por usuario típico en el sitio.

Sitio Infobae: El sitio Infobae tiene 205.882 visitas por día que en promedio pasan unos 2 minutos por usuario típico en el sitio.

¿Es difícil de realizar una distribución? En absoluto, si se cuenta con la serie que indica el tiempo que estuvo cada una de las visitas en la página entonces lo que se hace es lo siguiente: En Excel dentro de Herramientas existe una opción que dice “Análisis de Datos” donde allí se puede elegir “Histograma”. Esta opción nos permite graficar automáticamente la distribución deseada. E incluso elegir rangos de corte deseados. Por ejemplo:



OTRAS CONSIDERACIONES:

- α Considero que para realizar análisis más elaborados es necesario pasar de utilizar un único indicador a utilizar varios. De lo contrario siempre serán análisis unidimensionales. En casos simples esto no tiene problema alguno aunque para otros análisis si serían necesarios. Para empezar me parece bien usar lo simple.
- α Una análisis interesante, aunque habría que hacerlo con sumo cuidado, sería el de trazar en el mismo gráfico de distribución la vertical del “tiempo de lectura estimado”. Visualmente:



De esa manera podríamos ver que mientras que el tiempo promedio de visita del 68% del total de los visitantes se encuentra en, por ejemplo, 4 minutos, el “tiempo de lectura estimado” se encuentra en 8 minutos. Es decir, únicamente el 10% (por arrojar un número) se queda junto a la lectura todo el tiempo para el cual fue diseñada.

α **Otros indicadores posibles:**

a) Total Visitas x Tiempo Promedio del total de visitas

En este caso el indicador diría el tiempo total que fue utilizado el sitio.

b) Total Visitas / Tiempo Total de Utilización del Sitio

En este caso nos diría “visitas promedio por minuto” (en el caso en que estuviera medido en minutos). Sería una medida de la rotación de gente que posee el sitio. Es de suponer que distintas horas del día tengan distintas rotaciones.

c) (Tiempo total de Utilización del Sitio) / (Tiempo Total de Lectura del Sitio)

Este indicador nos podría decir el porcentaje de utilización del sitio en función del tiempo para el cual fue diseñado.

PROTOTIPOS

¿Qué sería interesante? Realizar un par de ensayos con algunos sitios. De esa manera podríamos ver algunos casos concretos y sacar algunas conclusiones más empíricas.