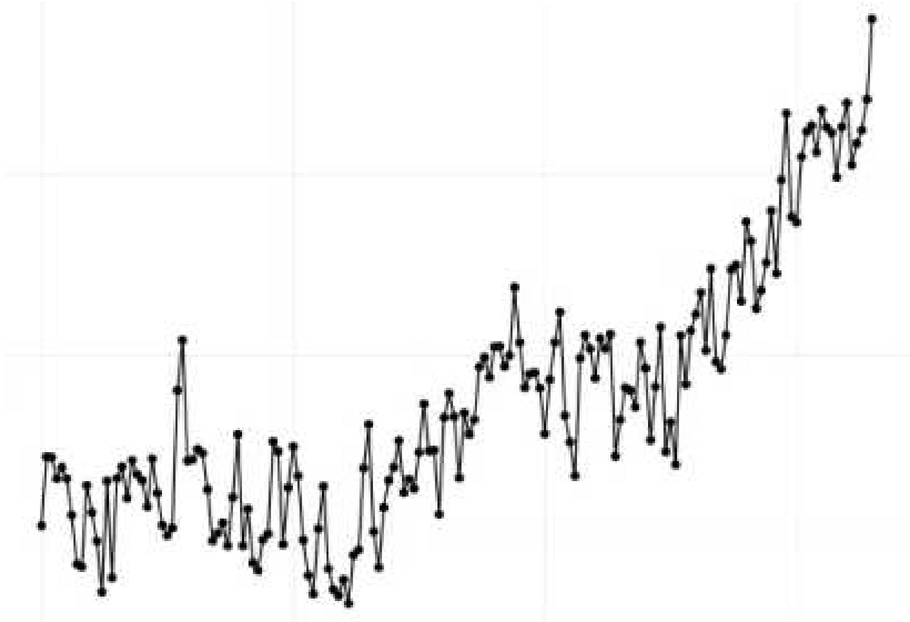
Amontonamiento - principio estadístico

Referencia: Mary L. Boas, Mathematical Methods in the Physical Sciences 732–736



A un cierto tiempo en un sismograma, una medición, X, de una señal sísmica, tiene ruido incorporado. Podemos decir que la medición tiene una función de densidad de población $f(x)$, donde el valor esperado de X es μ (el promedio de $f(x)$), y la varianza en X es σ^2. Si hay mayor ruido en la señal, su desviación estándar σ sería mayor.

- μ se llama el promedio de la población $f(x)$
- σ^2 se llama la varianza en la población $f(x)$

Ahora vamos a tomar n mediciones independientes de X: es decir que tenemos X_i con i de 1 a n. Se puede pensar en las diferentes mediciones como sismogramas de diferentes eventos tomados en diferentes estaciones, pero con la distancia fuente-estación igual; o en la exploración con fuentes activas se puede generar la misma fuente varias veces y registrar las ondas que genera. Ahora vamos a promediar estas n mediciones:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

Las siguientes preguntas requieren conocimiento básico sobre valores esperados y varianzas, que se presenta en la apéndice donde se introduce cuatro ecuaciones (A1 - A4) que sirven para el análisis.

¿Cuál es el valor esperado de $\bar{X}$?

Para calcular el valor esperado del promedio de las n mediciones, $\bar{X}$, consideremos

$$\begin{aligned}
E[\bar{X}] &= E\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] \\
&= \frac{1}{n}E\left[\sum_{i=1}^{n} X_i\right] && \text{usando la relación (A1)} \\
&= \frac{1}{n}\sum_{i=1}^{n} E[X_i] && \text{usando la relación (A2)} \\
&= \frac{1}{n}\sum_{i=1}^{n} \mu = \mu
\end{aligned}$$

Entonces el valor esperado del promedio $\bar{X}$ es la misma que el valor esperado de cada medición. Eso no es una gran sorpresa.

¿Cuál es la varianza en $\bar{X}$?

La varianza en $\bar{X}$ considere la diferencia entre el promedio $\bar{X}$ y su valor esperado $E[\bar{X}]$:

$$Var(\bar{X}) = E\left[\left(\bar{X} - E[\bar{X}]\right)^2\right]$$

Para calcular su valor relativo a la varianza en una medición (σ^2), consideremos

$$\begin{aligned}
Var(\bar{X}) &= Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \\
&= \frac{1}{n^2}Var\left(\sum_{i=1}^{n} X_i\right) && \text{usando la relación (A3)} \\
&= \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) && \text{usando la relación (A4)} \\
&= \frac{1}{n^2}\sum_{i=1}^{n} \sigma^2 = \frac{\sigma^2}{n}
\end{aligned}$$

Entonces $Var(\bar{X})$, llamado la varianza en la muestra, se reduce por un factor de $1/n$ comparado con la varianza en la población de una sola medición.

En conclusión, $\bar{X}$, el promedio de las mediciones, es una estimación de μ que aumenta en precisión cuando aumentamos la cantidad, n, de mediciones. Por consecuencia, la desviación estándar del promedio de las n mediciones se reduce por un factor de $1/\sqrt{n}$.

En términos prácticos, como ejemplo, el promedio de 100 mediciones independientes de una señal sísmica aumentaría la taza señal:ruido de las mediciones individuales por un factor de 10. Eso es el poder del principio de amontonamiento.

Apéndice

Valor esperado

El valor esperado, o promedio, de una medición X podemos escribir por $E[X] = \mu$. El valor esperado satisface las siguientes propiedades:

- El valor esperado de una medición X multiplicado por una constante es igual a la constante multiplicado por el valor esperado de la medición.

$$E[kX] = kE[X] \tag{A1}$$

- El valor esperado de la suma de mediciones X y Y es igual a la suma de sus valores esperados.

$$E[X + Y] = E[X] + E[Y] \tag{A2}$$

Varianza

La propiedad $X - E[X]$ es una variable aleatoria que mide la desviación de la medición X de su valor esperado. En otras palabras, $X - E[X]$ es indicativa del ruido que contiene un punto de la serie de tiempo. Podemos definir la varianza en la medición:

$$\begin{aligned} Var(X) &= E\left[(X - E[X])^2\right] \\ \sigma^2 &= E\left[(X - \mu)^2\right] \end{aligned}$$

Ahora consideremos la varianza en la medición X multiplicado por una constante, y manipular usando la ecuación A1:

$$\begin{aligned} Var(kX) &= E\left[(kX - E[kX])^2\right] \\ &= E\left[(kX - kE[X])^2\right] \\ &= E\left[k^2 (X - E[X])^2\right] \\ &= k^2 E\left[(X - E[X])^2\right] \\ \Rightarrow Var(kX) &= k^2 Var(X) \end{aligned} \tag{A3}$$

Y, últimamente veremos la varianza en la suma de mediciones X y Y, usando las ecuaciones A1 y A2:

$$\begin{aligned} Var(X + Y) &= E\left[((X + Y) - E[X + Y])^2\right] \\ &= E\left[(X + Y)^2 - 2(X + Y)E[X + Y] + E[X + Y]^2\right] \\ &= E\left[(X + Y)^2 - 2(X + Y)(E[X] + E[Y]) + (E[X] + E[Y])^2\right] \\ &= E\left[X^2 + 2XY + Y^2 - 2XE[X] - 2YE[Y] - 2XE[Y] - 2YE[X] \right. \\ &\quad \left. + E[X]^2 + E[Y]^2 + 2E[X]E[Y]\right] \\ &= E\left[(X^2 - 2XE[X] + E[X]^2) + (Y^2 - 2YE[Y] + E[Y]^2) \right. \\ &\quad \left. + 2(XY - XE[Y] - YE[X] + E[X]E[Y])\right] \\ &= E\left[(X - E[X])^2\right] + E\left[(Y - E[Y])^2\right] + 2E\left[(X - E[X])(Y - E[Y])\right] \end{aligned}$$

Llegando a la relación

$$Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$$

La covarianza entre mediciones X y Y, es decir $Cov(X,Y) = E\left[(X - E[X])(Y - E[Y])\right]$, es el valor esperado del producto entre

i la desviación de X de su valor esperado, y

ii la desviación de Y de su valor esperado.

Si X y Y están mediciones independientes, por ejemplo el ruido en la medición X tiene ningún relación con el ruido en la medición Y, la correlación entre las dos mediciones es cero, y su covarianza es cero. En este caso:

$$Var(X+Y) = Var(X) + Var(Y) \qquad (X,\ Y \text{ independientes entre ellas}) \tag{A4}$$