Artificial Neural Networks

Lecture 11: Error Propagation

2022-23

Propagation of Errors; Errors When Changing Variables

- Within the subject of Statistical Machine Learning we will explore quantities x that are distributed, i.e. there exists $\mathcal{P} : \mathbb{D} \to \mathbb{R}^+ \cup \{0\}$ and $\sum_{x \in \mathbb{D}} \mathcal{P}(x) = 1$ that describes the statistics of x.
- Every measurement process produces an estimate $\overline{x}$ (for a quantity of interest x) and an error $\overline{\sigma}_x$ which provides an estimate of the statistical dispersion around the estimate.
- Both quantities are estimates. They are not the true values of the $x_0 = \sum_{x \in \mathbb{D}} \mathcal{P}(x) x = \mathbb{E}[x]$ and $\sigma_x^2 = \sum_{x \in \mathbb{D}} \mathcal{P}(x)(x - x_0)^2 = \mathbb{E}[(x - \mathbb{E}[x])^2]$. These values, x_0 and σ_x, are, in general, not accessible but can be estimated.

Propagation of Errors; Errors When Changing Variables

- Suppose we measure x which has a mean value x_0 and a variance σ_x^2 and y with a mean y_0 and variance σ_y^2.
- Suppose we have a function $G(x,y)$ and wish to determine the variance of $G(x,y)$,i.e., propagate the errors in x and y to G. Thus

$$
\begin{aligned}
G(x,y) &= G(x_0 + x - x_0, y_0 + y - y_0) \\
&= G(x_0,y_0) + \left.\frac{\partial G}{\partial x}\right|_{x_0,y_0}(x - x_0) + \left.\frac{\partial G}{\partial y}\right|_{x_0,y_0}(y - y_0) + O(\Delta^2) \\
G_0 &= G(x_0,y_0) + O(\Delta^2) \\
\sigma_G^2 &= \left(\left.\frac{\partial G}{\partial x}\right|_{x_0,y_0}\right)^2 \sigma_x^2 + \left(\left.\frac{\partial G}{\partial y}\right|_{x_0,y_0}\right)^2 \sigma_y^2 + O(\Delta^4).
\end{aligned}
$$

.

Worked Problem

- Suppose we take n independent measurements of the same quantity x. Suppose each measurement x_i has the same mean $x_0$and variance σ_x^2.
- Given the following definition

$$G(\{x_i\}) = \frac{1}{n}\sum_{i=1}^{n} x_i,$$

find the mean and variance of G.

Solution

- The mean is given by (up to corrections of $O(\Delta^2)$)

$$G_0 = \frac{1}{n}\sum_{i=1}^{n} x_0 = x_0 .$$

- The variance is given by (up to corrections of $O(\Delta^4)$)

$$\sigma_G^2 = \sum_{i=1}^{n}\left(\left.\frac{\partial G}{\partial x_i}\right|_{x_0}\right)^2 \sigma_x^2 = \sum_{i=1}^{n}\frac{1}{n^2}\sigma_x^2 = \frac{\sigma_x^2}{n}.$$

Interpretation

- The mean $x_0$is in general inaccessible. We usually substitute $x_0$with the arithmetic mean $\frac{1}{n}\sum_{i=1}^{n} x_i \equiv \overline{x}$.
- Let us define the experimental variances

$$\overline{\sigma^2_{\mathrm{exp}}} \equiv \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2$$
$$\sigma^2_{\mathrm{exp}} \equiv \mathbb{E}[(x - \overline{x})^2].$$

- The first expression can be measured, the second is, in general, inaccessible (we do not know the distribution $\mathcal{P}(x)$).

Interpretation

- The experimental variance is linked to the true variance in the following way:

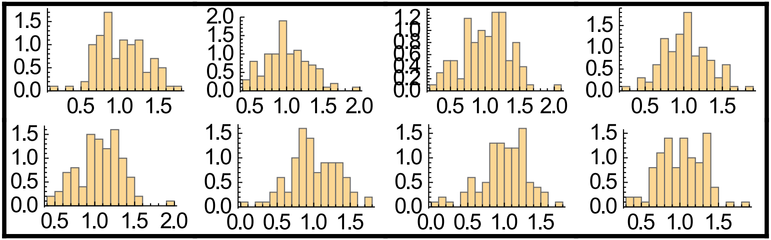
$$
\begin{aligned}
\sigma^2_{\mathrm{exp}} &= \mathbb{E}[(x-\overline{x})^2] = \mathbb{E}[(x_i-\overline{x})^2] \\
&= \mathbb{E}\left[\left(\frac{n-1}{n}x_i - \frac{1}{n}\sum_{j\neq i} x_j\right)^2\right] \\
&= \left(\frac{n-1}{n}\right)^2 \mathbb{E}[x_i^2] - 2\frac{n-1}{n^2}\mathbb{E}[x_i]\sum_{j\neq i}\mathbb{E}[x_j] + \frac{1}{n^2}\mathbb{E}\left[\sum_{j\neq i} x_j^2 + 2\sum_{j\neq i}\sum_{k\neq i,j} x_j x_k\right] \\
&= \left(\frac{n-1}{n}\right)^2 \mathbb{E}[x_i^2] - 2\frac{n-1}{n^2}\mathbb{E}[x_i]\sum_{j\neq i}\mathbb{E}[x_j] + \frac{1}{n^2}\sum_{j\neq i}\mathbb{E}\left[x_j^2\right] + \frac{2}{n^2}\sum_{j\neq i}\sum_{k\neq i,j}\mathbb{E}[x_j]\mathbb{E}[x_k] \\
&= \left(\frac{n-1}{n}\right)^2 \mathbb{E}[x^2] - 2\left(\frac{n-1}{n}\right)^2\mathbb{E}[x]^2 + \frac{n-1}{n^2}\mathbb{E}[x^2] + \frac{2}{n^2}\frac{(n-1)^2-(n-1)}{2}\mathbb{E}[x]^2 \\
&= \frac{n-1}{n^2}(n-1+1)\mathbb{E}[x^2] - 2\frac{n-1}{n^2}\left(n-1-\frac{n-1-1}{2}\right)\mathbb{E}[x]^2 = \frac{n-1}{n}\left(\mathbb{E}[x^2]-\mathbb{E}[x]^2\right) \\
&= \frac{n-1}{n}\sigma_x^2
\end{aligned}
$$

Interpretation

- If we estimate $\sigma^2_{\exp}$ using $\overline{\sigma^2_{\exp}}$ we can estimate the true variance by
$$\sigma_x^2 \approx \frac{n}{n-1}\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2.$$
- Therefore

$$G_0 \approx \frac{1}{n}\sum_{i=1}^{n} x_i = \overline{x}$$
$$\sigma_G^2 \approx \frac{1}{n(n-1)}\sum_{i=1}^{n}(x_i - \overline{x})^2.$$

Interpretation



Interpretation