Probability vs. Probability Density Function (PDF)
機率分佈 與 機率密度函數

Table 1 為一次考試的結果，考 80、70、30 分的學生人數分別有 6、12、2 人。因為總人數 6+12+2=20 人，因此這樣的結果，也可以用機率的方式呈現，列在 Table 2 。

Table 1. 考試結果成績分布

i	x_i 成績	F_i 人數
1	80	6
2	70	12
3	30	2

Table 2. 考試結果機率分佈

i	x_i 成績（亂數）	F_i 機率
1	80	30%
2	70	60%
3	30	10%

Table 3 與 Table 4，分別利用 Table 1 and Table 2 所提供的資料計算這次考試成績的平均值、變異數、標準差。

Table 3. 利用 Table 1 所提供的資料計算成績的平均值、變異數、標準差

總人數	$\sum_{i=1}^{3} F_i = F_1 + F_2 + F_3 = 20$
成績平均值 (平均分數)	$\bar{x} = \dfrac{80\times 6 + 70\times 12 + 30\times 2}{20}$ or $\bar{x} = \dfrac{\sum_{i=1}^{3} x_i F_i}{\sum_{i=1}^{3} F_i}$
成績變異數	$\text{Variance} = \sigma^2 = \dfrac{(80-\bar{x})^2\times 6 + (70-\bar{x})^2\times 12 + (30-\bar{x})^2\times 2}{20}$ or $\text{Variance} = \sigma^2 = \dfrac{\sum_{i=1}^{3} (x_i-\bar{x})^2 F_i}{\sum_{i=1}^{3} F_i}$
成績標準差	Standard deviation $\sigma = \sqrt{\text{Variance}}$

Table 4. 利用 Table 2 所提供的資料計算成績的平均值、變異數、標準差

機率和	$\sum_{i=1}^{3} F_i = F_1 + F_2 + F_3 = 0.3 + 0.6 + 0.1 = 1$
成績平均值 (平均分數)	$\bar{x} = 80\times0.3 + 70\times0.6 + 30\times0.1$ or $\bar{x} = \sum_{i=1}^{3} x_i F_i$ （因為 $\sum_{i=1}^{3} F_i = 1$ ）
成績變異數	$\text{Variance} = \sigma^2 = (80-\bar{x})^2\times0.3 + (70-\bar{x})^2\times0.6 + (30-\bar{x})^2\times0.1$ or $\text{Variance} = \sigma^2 = \sum_{i=1}^{3}(x_i-\bar{x})^2 F_i$ （因為 $\sum_{i=1}^{3} F_i = 1$ ）
成績標準差	Standard deviation $\sigma = \sqrt{\text{Variance}}$

Q1:
利用 Table 2 的資料，求平均值與變異數時，部份同學的答案如下。
知道錯哪裡嗎？

成績平均值 $\bar{x} = \dfrac{80\times0.3 + 70\times0.6 + 30\times0.1}{3}$

成績變異數 $\text{Variance} = \dfrac{(80-\bar{x})^2\times0.3 + (70-\bar{x})^2\times0.6 + (30-\bar{x})^2\times0.1}{3}$

Hint: 想想看 30% 與 6 人以及總人數 20 人之間的關係。

Let us consider a continuous probability density function $f(x)$. Let the probability of $(x_i - \frac{\Delta x}{2}) < x < (x_i + \frac{\Delta x}{2})$ is F_i, where $x_{i+1} = x_i + \Delta x$. It yields $F_i = \int_{x_i - \frac{\Delta x}{2}}^{x_i + \frac{\Delta x}{2}} f(x)dx$

由以上說明（英文），我們可以將「連續變化的機率密度函數 probability density function $f(x)$」與「一個個亂數的機率分佈」之間的關係，列表說明於 Table 5。進一步，我們也可以參考 Table 3 and Table 4 的格式，寫出「滿足機率密度函數 probability density function $f(x)$ 之亂數群」的變異數與標準差的計算方式，如 Table 6 所示。

Table 5. 由機率密度函數估算等間距亂數的發生機率

i	x_i	F_i
		$\int_{-\infty}^{x_1 - \frac{\Delta x}{2}} f(x)dx \to 0$
1	x_1	$F_1 = \int_{x_1 - \frac{\Delta x}{2}}^{x_1 + \frac{\Delta x}{2}} f(x)dx$
2	$x_2 = x_1 + \Delta x$	$F_2 = \int_{x_2 - \frac{\Delta x}{2}}^{x_2 + \frac{\Delta x}{2}} f(x)dx$
…	**…**	…
N	$x_N = x_{N-1} + \Delta x$	$F_N = \int_{x_N - \frac{\Delta x}{2}}^{x_N + \frac{\Delta x}{2}} f(x)dx$
		$\int_{x_N + \frac{\Delta x}{2}}^{\infty} f(x)dx \to 0$

Table 6. 利用 Table 5 所提供的資料計算成績的平均值、變異數、標準差

Mean 平均值	$\bar{x} = \frac{\sum_{i=1}^{N} x_i F_i}{\sum_{i=1}^{N} F_i} = \frac{\int_{-\infty}^{\infty} x f(x)dx}{\int_{-\infty}^{\infty} f(x)dx}$ If $\int_{-\infty}^{\infty} f(x)dx = 1$ then $\bar{x} = \int_{-\infty}^{\infty} x f(x)dx$
Varance 變異數	$\text{Variance} = \sigma^2 = \frac{\sum_{i=1}^{N} (x_i - \bar{x})^2 F_i}{\sum_{i=1}^{N} F_i} = \frac{\int_{-\infty}^{\infty} (x - \bar{x})^2 f(x)dx}{\int_{-\infty}^{\infty} f(x)dx}$ If $\int_{-\infty}^{\infty} f(x)dx = 1$ then $\text{Varance} = \sigma^2 = \int_{-\infty}^{\infty} (x - \bar{x})^2 f(x)dx$
標準差	Standard deviation $\sigma = \sqrt{\text{Variance}}$