6 Mean and Variance

The **mean** μ or **expected value** $E[X]$ of the discrete random variable is known as the center of a probability distribution

$$\mu = E[X] = \sum_x x f(x) \tag{1}$$

The **variance** σ^2 or $V[X]$ of X is a measure of the dispersion, or variability in the distribution

$$\sigma^2 = V[X] = E(X-\mu)^2 = \sum_x (x-\mu)^2 f(x) = \sum_x x^2 f(x) - \mu^2 \tag{2}$$

Lastly the standard deviation σ of a distribution is

$$\sigma = +\sqrt{\sigma^2} = +\sqrt{V[X]} \tag{3}$$

At first glance variance may look daunting to calculate, but all equations are the same.

Proof. Show that $E(X-\mu)^2 = \sum_x (x-\mu)^2 f(x) = \sum_x x^2 f(x) - \mu^2$

$$E(X-\mu)^2 = \sum_x (X-\mu)^2 f(x)$$

and

$$\begin{aligned}
\sum_x (X-\mu)^2 f(x) &= \sum_x (x^2 - 2x\mu + \mu^2) f(x) \\
&= \sum_x x^2 f(x) - 2\big(\sum_x x f(x)\big)\mu + \sum_x \mu^2 f(x) \\
&= \sum_x x^2 f(x) - 2\mu^2 + \mu^2 \sum_x f(x) \\
&= \sum_x x^2 f(x) - 2\mu^2 + \mu^2(1) \\
&= \sum_x x^2 f(x) - \mu^2
\end{aligned}$$

Use case 1: μ **:** Principal of Least Squares: Find the value a that minimizes the sum of the squares

$$L(a) = \sum_x (x-a)^2 f(x)$$

In order to minimize $L(a)$, we need to find $dL/da = 0$

$$\frac{d}{da} \sum_x (x-a)^2 f(x) = \sum_x -2(x-a) f(x) = 0$$

$$\sum_x x f(x) = a \sum_x f(x)$$

$$\mu = a$$

So we can see the value of $a = \mu$ minimizes $L(a)$

Use case 2: μ **:** Center of Mass: Suppose we have masses $f(x_i)$ at locations x_i on a massless wire. Let m be the center of mass defined by

$$\sum_x f(x)(x-m) = 0$$

Since m is the center of a function which distributes mass, we know $m = \mu$ because μ describes the center of a distribution.

Use case 3: μ **:** μ is the average value of X over repeated observations.

Use case 4: μ **:** For equally likely outcomes μ is the average of all possible outcomes.

6.1 Expected Value of a Function of X

Let $Y = h(X)$ so it is also a random variable with its own p.m.f $f_Y(y)$. By definition

$$E[Y] = \sum_y y f_Y(y) \tag{4}$$

It can be shown that the equation above can be written as a function of X

$$E[h(X)] = h(x)f_X(x) \tag{5}$$

So we notice that $V[X] = E[(X-\mu)^2]$, we can also find that $V[x] = E[X^2] - (E[X])^2$

Proof:

$$\begin{aligned} E[(X-\mu)^2] &= E[X^2 - 2X\mu + \mu^2] \\ &= E[X^2] - 2\mu E[X] + \mu^2 \\ &= E[X^2] - \mu^2 \\ &= E[X^2] - (E[x])^2 \end{aligned}$$

The following shows the linearity of mean and variance: Let a and b be constants

1. $E[aX+b] = aE[X] + b$

   Proof:

$$\begin{aligned} E[aX+b] &= \sum_x (ax+b)f(x) \\ &= a\sum_x xf(x) + b\sum_x f(x) \\ &= aE[X] + b \end{aligned}$$

2. $V[aX+b] = a^2V[X]$

   Proof:

$$\begin{aligned} V[aX+b] &= E[(aX+b)^2] - (E[aX+b])^2 \\ &= E[(aX)^2 + 2aXb + b^2] - (aE[X]+b)^2 \\ &= a^2E[X^2] + 2abE[X] + b^2 - a^2(E[X])^2 - 2abE[X] + b^2 \\ &= a^2E[X^2] - a^2(E[X])^2 \\ &= a^2(E[X^2] - (E[X])^2) \\ &= a^2V[X] \end{aligned}$$