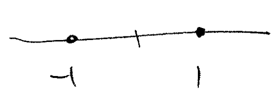
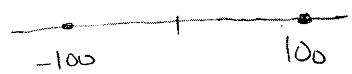
2.3 Variance and Standard Deviation of Discrete Random Variable

Say: Recall that we looked at two features of a set of observations: its center and its spread (variation). The center is given by the mean, while the spread (variation) is measured by the variance. The square root of the variance is called the standard deviation, and it also measures the spread, but with better units.



and



both have mean 0, but variance of the latter is much more.

We similarly have notions of variance and standard deviation for random variables.

Def Let $E[Y] =: \mu$. Then the variance of Y is

$$V(Y) := E[(Y-\mu)^2].$$

The standard deviation of Y is

$$SD(Y) := \sqrt{V(Y)}$$

Thm $V(Y) = E[Y^2] - (E[Y])^2$

Ex The variance in the previous example is

$$V(Y) = E[Y^2] - (E[Y])^2$$

$$= \frac{27}{5} - \left(\frac{11}{5}\right)^2 = \underbrace{\frac{135}{25} - \frac{121}{25}}_{\text{skip}} = \frac{14}{25}$$

$$SD(Y) = \sqrt{V(Y)} = \sqrt{\frac{14}{25}} = \frac{\sqrt{14}}{5}.$$

<u>say</u>: when you shift data points left or right uniformly, the spread doesn't change (though the center will change). When you multiply all data points by a factor it multiplies the variance by the square of the factor.

<u>Rem</u> $V(-)$ is not a linear operator.

<u>Thm</u> If a and b are real constants, then $V(aY+b) = a^2 V(Y)$.

<u>Ex</u> In the prev. ex., $V(2Y) = 2^2 V(Y) = 4 \cdot \frac{14}{25} = \frac{56}{25}$

$$V(Y+10) = V(Y) = \frac{14}{25}$$

$$V(2Y+10) = 2^2 V(Y) = \frac{56}{25}.$$