

Why is the Variance of the Sum of Two Independent Random Variables the Sum of the Variances?

Imagine two such random variables X and Y .

| X | probability |
|----------|--------------------|
| x_1 | p_1 |
| x_2 | p_2 |
| \dots | \dots |
| x_n | p_n |

| Y | probability |
|----------|--------------------|
| y_1 | q_1 |
| y_2 | q_2 |
| \dots | \dots |
| y_m | q_m |

Since X and Y are independent random variables, the probability of X taking on the value x_i and Y the value y_j is simply the product $p_i q_j$.

Below is the table describing the random variable $X + Y$. Some values may appear more than once in the nm rows below, but this does not throw off our formulae for the variance.

| X + Y | probability |
|--------------|--------------------|
| $x_1 + y_1$ | $p_1 q_1$ |
| $x_1 + y_2$ | $p_1 q_2$ |
| \dots | \dots |
| $x_n + y_m$ | $p_n q_m$ |

Now

$$\begin{aligned}
 \text{Var}(X + Y) &= \sum_{i,j} p_i q_j (x_i + y_j - \mu_X - \mu_Y)^2 \\
 &= \sum_{i,j} p_i q_j (x_i - \mu_X)^2 + \sum_{i,j} p_i q_j (y_j - \mu_Y)^2 + 2 \sum_{i,j} p_i q_j (x_i - \mu_X)(y_j - \mu_Y) \\
 &= (\sum_j q_j) \text{Var}(X) + (\sum_i p_i) \text{Var}(Y) + 2 (\sum_i p_i (x_i - \mu_X)) (\sum_j q_j (y_j - \mu_Y)) \\
 &= 1 \cdot \text{Var}(X) + 1 \cdot \text{Var}(Y) + 2 \cdot 0 \cdot 0 \\
 &= \text{Var}(X) + \text{Var}(Y).
 \end{aligned}$$