Properties of variance:

1. It is not true in general that \( \text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y) \).

2. \( \text{Var}(aX) = a^2 \cdot \text{Var}(X) \), so \( \sigma_{aX} = a \cdot \sigma_X \).

3. \( \text{Var}(X+b) = \text{Var}(X) \), so \( \sigma_{X+b} = \sigma_X \).

Proof of (3):

\[
\text{Var}(X+b) = E \left[ (X+b) - E(X+b) \right]^2 = E \left[ (X + b - E[X] - b)^2 \right] \quad (\text{by props of expectation})
\]

\[
= E \left[ (X - E[X])^2 \right] = \text{Var}(X). \quad \text{QED.}
\]

Ex. (The matching problem)

\( N \) homeworks are returned to \( N \) students at random. How many students will get their own homework on average?

(Recall that we calculated \( P \{ \text{at least } 1 \text{ student gets his/her own HW} \} = 1 - \frac{1}{e} \), see p. 17.)

Method of indicators: \( X = \# \text{ students getting their own HW} \).

Represent

\[
X = \sum_{i=1}^{N} X_i \quad \text{where} \quad X_i = \begin{cases} 
1, & \text{if student } i \text{ gets his/her own HW} \\
0, & \text{otherwise} 
\end{cases}
\]

\[
\Rightarrow E[X] = \sum_{i=1}^{N} E[X_i], \quad ?
\]

\[
E[X_i] = 1 \cdot P[X_i = 1] + 0 \cdot P[X_i = 0] = P[X_i = 1] = P\{ \text{student } i \text{ gets own HW} \} = \frac{1}{n}. \quad \text{(why?)}
\]

Therefore

\[
E[X] = N \cdot \frac{1}{N} = 1
\]

Ex. What is the standard deviation of the students who get their own HW?

\[
E[X^2] = E \left[ \sum_{i=1}^{N} X_i \right]^2 = E \left[ \sum_{i=1}^{N} X_i^2 + \sum_{i \neq j} X_i X_j \right] = \sum_{i=1}^{N} E[X_i^2] + \sum_{i \neq j} E[X_i X_j].
\]

\[
E[X_i^2] = E[X_i] = \frac{1}{N}, \quad \text{as before.}
\]

\[
E[X_i X_j] = 1 \cdot P[X_i, X_j = 1] + 0 \cdot P[X_i, X_j = 0] = P\{ X_i = 1 \text{ and } X_j = 1 \} = P\{ \text{both students } i, j \text{ get their own HW} \} = \frac{1}{N(N-1)}. \quad \text{(Why? e.g. by conditioning on one of students)}
\]

Therefore

\[
E[X^2] = N \cdot \frac{1}{N} + N(N-1) \cdot \frac{1}{N(N-1)} = 2.
\]

}\[
\text{Hence } \text{Var}(X) = E[X^2] - (E[X])^2 = 2 - 1 = 1, \quad \sigma_X = \sqrt{1} = 1.
\]

36
4.6. Bernoulli and Binomial distributions

**Bernoulli**

- Experiment has two outcomes: success, failure. \( P(\text{success}) = p, \quad P(\text{failure}) = 1 - p \)
- Let r.v. \( X \) be the indicator of success, hence \( X \) takes two values: 1 with prob. \( p \), 0 with prob. \( 1 - p \)

Formally:

**Def.** A Bernoulli r.v. with parameter \( p \) is a r.v. \( X \) with pmf

- \( p(1) = p, \quad p(0) = 1 - p \)
- Notation: \( X \sim Bernoulli(p) \)

- \( E(X) = 1 \cdot p + 0 \cdot (1 - p) = p \)
- \( E[X^2] = E[X] = p \); \quad \text{Var}(X) = p - p^2 = p(1 - p) \)

**Binomial**

- Experiment consists of \( n \) independent trials, each having success with prob. \( p \), failure with prob. \( 1 - p \).
- Let r.v. \( X \) be the \# of successes

\[ p(k) = P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \]

Formally:

**Def.** A Binomial r.v. with parameters \( n, p \) is a r.v. \( X \) with pmf

- \( p(k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \ldots, n \)
- Notation: \( X \sim Binomial(n, p) \)

-
Jack hits target 70% of time. What is the prob. that he hits target in at least 8 of 10 shots?

\[ X = \text{# of hits} \, \quad X \sim \text{Binom}(10, 0.7) \]

\[ P\{X \geq 8\} = 1 - P(9) - P(10) = \binom{10}{9} 0.7^9 0.3^1 + \binom{10}{10} 0.7^{10} 0.3^0 = 0.93 \]

- let \( X \sim \text{Binom}(n, p) \). One can represent

\[ X = \sum_{i=1}^{n} X_i \quad \text{where} \quad X_i = \begin{cases} 1, & \text{success in trial } i \\ 0, & \text{otherwise} \end{cases} \]

\[ X_i \sim \text{Bernoulli}(p) \]

- \( E[X] = \sum_{i=1}^{N} E[X_i] = np \).

- \( E[X^2] = E[\left( \sum_{i=1}^{N} X_i \right)^2] = \sum_{i=1}^{N} E[X_i^2] + \sum_{i \neq j} E[X_i X_j] \)

\[ = np + np(n-1)p^2 \quad \text{(by independence)} \]

\[ \Rightarrow \text{Var}(X) = np + np(n-1)p^2 - (np)^2 = np(1-p) \]

- Coin tossed \( n \) times. \( X = \# \text{ heads} \, \quad X \sim \text{Binom}(n, \frac{1}{2}) \)

\[ E[X] = \frac{n}{2} \Rightarrow \text{Var}(X) = \frac{n}{4} \Rightarrow \sigma_X = \frac{\sqrt{n}}{2} \]

So, typically there are \( \frac{n \pm \sqrt{n}}{2} \) heads.

E.g. for \( n = 10^6 \Rightarrow 500,000 \pm 500 \) heads.
Ex (overbooking)

Flight A: 10 tickets sold, the plane has 9 seats.
Flight B: 20 tickets sold, the plane has 18 seats.

Passengers show up with prob. 0.9 each.
Which flight is more likely to get overbooked?

\[ X_A = \# \text{ passengers who show up for flight A} \]
\[ X_B = \# \text{ passengers who show up for flight B} \]

\[ X_A \sim \text{Binomial}(10, 0.9), \quad X_B \sim \text{Binomial}(20, 0.9) \]

\[ P(\text{A overbooked}) = P\{X_A > 9\} = P\{X_A = 10\} = \binom{10}{10} 0.9^{10} \approx 0.346 \]

\[ P(\text{B overbooked}) = P\{X_B > 18\} = P\{X_B = 19\} + P\{X_B = 20\} = \binom{20}{19} 0.9^{19} \cdot 0.1 + \binom{20}{20} 0.9^{20} = 0.388 \]

Answer: B