

Math 461 Fall 2021

Renming Song

University of Illinois at Urbana-Champaign

November 01, 2021

Outline

Outline

- 1 **General Info**
- 2 7.4 Covariance, variance of sums and correlations

HW9 is due Friday, 11/05, before the end of class. You can submit your HW9 via the course Moodle page. Make sure that your HW is uploaded successfully

Solution to HW8 is on my homepage.

HW9 is due Friday, 11/05, before the end of class. You can submit your HW9 via the course Moodle page. Make sure that your HW is uploaded successfully

Solution to HW8 is on my homepage.

Outline

- 1 General Info
- 2 7.4 Covariance, variance of sums and correlations**

Now we are going to use the formula

$$\begin{aligned}\text{Var}\left(\sum_{i=1}^n X_i\right) &= \sum_{i=1}^n \text{Var}(X_i) + \sum_{i \neq j} \text{Cov}(X_i, X_j) \\ &= \sum_{i=1}^n \text{Var}(X_i) + 2 \sum_{i < j} \text{Cov}(X_i, X_j).\end{aligned}$$

to find the variance of some complicated random variables.

Example 8

A group of N people throw their hats into the center of the room. The hats are mixed up, and each person randomly selects a hat. Let X be the number of people who get their own hats. Find $\text{Var}(X)$.

Now we are going to use the formula

$$\begin{aligned}\text{Var}\left(\sum_{i=1}^n X_i\right) &= \sum_{i=1}^n \text{Var}(X_i) + \sum_{i \neq j} \text{Cov}(X_i, X_j) \\ &= \sum_{i=1}^n \text{Var}(X_i) + 2 \sum_{i < j} \text{Cov}(X_i, X_j).\end{aligned}$$

to find the variance of some complicated random variables.

Example 8

A group of N people throw their hats into the center of the room. The hats are mixed up, and each person randomly selects a hat. Let X be the number of people who get their own hats. Find $\text{Var}(X)$.

For $i = 1, \dots, N$, let $X_i = 1$ if the i -th man gets his own hat and $X_i = 0$ otherwise. Then $X = X_1 + \dots + X_N$. Note that

$$E[X_i] = P(X_i = 1) = \frac{1}{N}, \quad \text{Var}(X_i) = \frac{1}{N} \left(1 - \frac{1}{N}\right).$$

Now let's find $\text{Cov}(X_i, X_j)$ for $i \neq j$. $X_i X_j$ is also a Bernoulli random variable.

$$E[X_i X_j] = P(X_i X_j = 1) = P(X_i = 1, X_j = 1) = \frac{1}{N(N-1)}.$$

Thus

$$\text{Cov}(X_i, X_j) = \frac{1}{N(N-1)} - \frac{1}{N^2}.$$

For $i = 1, \dots, N$, let $X_i = 1$ if the i -th man gets his own hat and $X_i = 0$ otherwise. Then $X = X_1 + \dots + X_N$. Note that

$$E[X_i] = P(X_i = 1) = \frac{1}{N}, \quad \text{Var}(X_i) = \frac{1}{N} \left(1 - \frac{1}{N}\right).$$

Now let's find $\text{Cov}(X_i, X_j)$ for $i \neq j$. $X_i X_j$ is also a Bernoulli random variable.

$$E[X_i X_j] = P(X_i X_j = 1) = P(X_i = 1, X_j = 1) = \frac{1}{N(N-1)}.$$

Thus

$$\text{Cov}(X_i, X_j) = \frac{1}{N(N-1)} - \frac{1}{N^2}.$$

$$\begin{aligned}\text{Var}(X) &= \text{Var}\left(\sum_{i=1}^N X_i\right) \\ &= \sum_{i=1}^N \text{Var}(X_i) + \sum_{i \neq j} \text{Cov}(X_i, X_j) \\ &= N \cdot \frac{1}{N} \left(1 - \frac{1}{N}\right) + N(N-1) \left(\frac{1}{N(N-1)} - \frac{1}{N^2}\right) \\ &= 1.\end{aligned}$$

Example 9

If n balls are randomly selected, without replacement, from a box containing N ($N > n$) balls, of which m are white. Let X be the number of white balls selected. Find $\text{Var}(X)$.

$$\begin{aligned}\text{Var}(X) &= \text{Var}\left(\sum_{i=1}^N X_i\right) \\ &= \sum_{i=1}^N \text{Var}(X_i) + \sum_{i \neq j} \text{Cov}(X_i, X_j) \\ &= N \cdot \frac{1}{N} \left(1 - \frac{1}{N}\right) + N(N-1) \left(\frac{1}{N(N-1)} - \frac{1}{N^2}\right) \\ &= 1.\end{aligned}$$

Example 9

If n balls are randomly selected, without replacement, from a box containing N ($N > n$) balls, of which m are white. Let X be the number of white balls selected. Find $\text{Var}(X)$.

For $i = 1, \dots, m$, let $Y_i = 1$ if the i -th white ball is among the selected. Then $X = Y_1 + \dots + Y_m$. Note that

$$E[Y_i] = \frac{\binom{N-1}{n-1}}{\binom{N}{n}} = \frac{n}{N}, \quad \text{Var}(Y_i) = \frac{n}{N} \left(1 - \frac{n}{N}\right).$$

Now let's find $\text{Cov}(Y_i, Y_j)$ for $i \neq j$. $Y_i Y_j$ is also a Bernoulli random variable.

$$E[Y_i Y_j] = P(Y_i = 1, Y_j = 1) = \frac{\binom{N-2}{n-2}}{\binom{N}{n}} = \frac{n(n-1)}{N(N-1)}.$$

Thus

$$\text{Cov}(Y_i, Y_j) = \frac{n(n-1)}{N(N-1)} - \frac{n^2}{N^2}.$$

For $i = 1, \dots, m$, let $Y_i = 1$ if the i -th white ball is among the selected. Then $X = Y_1 + \dots + Y_m$. Note that

$$E[Y_i] = \frac{\binom{N-1}{n-1}}{\binom{N}{n}} = \frac{n}{N}, \quad \text{Var}(Y_i) = \frac{n}{N} \left(1 - \frac{n}{N}\right).$$

Now let's find $\text{Cov}(Y_i, Y_j)$ for $i \neq j$. $Y_i Y_j$ is also a Bernoulli random variable.

$$E[Y_i Y_j] = P(Y_i = 1, Y_j = 1) = \frac{\binom{N-2}{n-2}}{\binom{N}{n}} = \frac{n(n-1)}{N(N-1)}.$$

Thus

$$\text{Cov}(Y_i, Y_j) = \frac{n(n-1)}{N(N-1)} - \frac{n^2}{N^2}.$$

$$\begin{aligned}\text{Var}(X) &= \text{Var}\left(\sum_{i=1}^m Y_i\right) \\ &= \sum_{i=1}^m \text{Var}(Y_i) + \sum_{i \neq j} \text{Cov}(Y_i, Y_j) \\ &= m \frac{n}{N} \left(1 - \frac{n}{N}\right) + m(m-1) \left(\frac{n(n-1)}{N(N-1)} - \frac{n^2}{N^2}\right).\end{aligned}$$

Example 10

n balls are randomly distributed into r boxes (so that each ball is equally likely to go to any of the r boxes). Let X be the number of empty boxes. Find $\text{Var}(X)$.

$$\begin{aligned}\text{Var}(X) &= \text{Var}\left(\sum_{i=1}^m Y_i\right) \\ &= \sum_{i=1}^m \text{Var}(Y_i) + \sum_{i \neq j} \text{Cov}(Y_i, Y_j) \\ &= m \frac{n}{N} \left(1 - \frac{n}{N}\right) + m(m-1) \left(\frac{n(n-1)}{N(N-1)} - \frac{n^2}{N^2}\right).\end{aligned}$$

Example 10

n balls are randomly distributed into r boxes (so that each ball is equally likely to go to any of the r boxes). Let X be the number of empty boxes. Find $\text{Var}(X)$.

For $i = 1, \dots, r$, let $X_i = 1$ if box number i is empty and $X_i = 0$ otherwise. Then $X = X_1 + \dots + X_r$. Note that for $i = 1, \dots, r$,

$$E[X_i] = P(X_i = 1) = \left(\frac{r-1}{r}\right)^n, \quad \text{Var}(X_i) = \left(\frac{r-1}{r}\right)^n \left(1 - \left(\frac{r-1}{r}\right)^n\right).$$

Now let's find $\text{Cov}(X_i, X_j)$ for $i \neq j$. $X_i X_j$ is also a Bernoulli random variable.

$$E[X_i X_j] = P(X_i = 1, X_j = 1) = \left(\frac{r-2}{r}\right)^n.$$

Thus

$$\text{Cov}(X_i, X_j) = \left(\frac{r-2}{r}\right)^n - \left(\frac{r-1}{r}\right)^{2n}.$$

For $i = 1, \dots, r$, let $X_i = 1$ if box number i is empty and $X_i = 0$ otherwise. Then $X = X_1 + \dots + X_r$. Note that for $i = 1, \dots, r$,

$$E[X_i] = P(X_i = 1) = \left(\frac{r-1}{r}\right)^n, \quad \text{Var}(X_i) = \left(\frac{r-1}{r}\right)^n \left(1 - \left(\frac{r-1}{r}\right)^n\right).$$

Now let's find $\text{Cov}(X_i, X_j)$ for $i \neq j$. $X_i X_j$ is also a Bernoulli random variable.

$$E[X_i X_j] = P(X_i = 1, X_j = 1) = \left(\frac{r-2}{r}\right)^n.$$

Thus

$$\text{Cov}(X_i, X_j) = \left(\frac{r-2}{r}\right)^n - \left(\frac{r-1}{r}\right)^{2n}.$$

$$\begin{aligned}
 \text{Var}(X) &= \text{Var}\left(\sum_{i=1}^r X_i\right) \\
 &= \sum_{i=1}^r \text{Var}(X_i) + \sum_{i \neq j} \text{Cov}(X_i, X_j) \\
 &= r \left(\frac{r-1}{r}\right)^n \left(1 - \left(\frac{r-1}{r}\right)^n\right) + r(r-1) \left(\left(\frac{r-2}{r}\right)^n - \left(\frac{r-1}{r}\right)^{2n}\right).
 \end{aligned}$$

Example 11

There are n types of coupons. Each newly obtained coupon is, independently, equally like to be any of the n types. Let X be the number of distinct types contained in a collection of k coupons. Find $\text{Var}(X)$.

$$\begin{aligned}\text{Var}(X) &= \text{Var}\left(\sum_{i=1}^r X_i\right) \\ &= \sum_{i=1}^r \text{Var}(X_i) + \sum_{i \neq j} \text{Cov}(X_i, X_j) \\ &= r \left(\frac{r-1}{r}\right)^n \left(1 - \left(\frac{r-1}{r}\right)^n\right) + r(r-1) \left(\left(\frac{r-2}{r}\right)^n - \left(\frac{r-1}{r}\right)^{2n}\right).\end{aligned}$$

Example 11

There are n types of coupons. Each newly obtained coupon is, independently, equally like to be any of the n types. Let X be the number of distinct types contained in a collection of k coupons. Find $\text{Var}(X)$.

For $i = 1, \dots, n$, let $X_i = 1$ if there is at least one type i coupon in the collection of k coupons and $X_i = 0$ otherwise. Then $X_1 + \dots + X_n$ is the number of distinct types in the collection of k coupons.

For $i = 1, \dots, n$,

$$P(X_i = 0) = \left(1 - \frac{1}{n}\right)^k, \quad P(X_i = 1) = 1 - \left(1 - \frac{1}{n}\right)^k.$$

For $i \neq j$,

$$\begin{aligned} P(X_i X_j = 0) &= P(X_i = 0) + P(X_j = 0) - P(X_i = 0, X_j = 0) \\ &= 2 \left(1 - \frac{1}{n}\right)^k - \left(1 - \frac{2}{n}\right)^k, \end{aligned}$$

$$P(X_i X_j = 1) = 1 - 2 \left(1 - \frac{1}{n}\right)^k + \left(1 - \frac{2}{n}\right)^k.$$

For $i = 1, \dots, n$, let $X_i = 1$ if there is at least one type i coupon in the collection of k coupons and $X_i = 0$ otherwise. Then $X_1 + \dots + X_n$ is the number of distinct types in the collection of k coupons.

For $i = 1, \dots, n$,

$$P(X_i = 0) = \left(1 - \frac{1}{n}\right)^k, \quad P(X_i = 1) = 1 - \left(1 - \frac{1}{n}\right)^k.$$

For $i \neq j$,

$$\begin{aligned} P(X_i X_j = 0) &= P(X_i = 0) + P(X_j = 0) - P(X_i = 0, X_j = 0) \\ &= 2 \left(1 - \frac{1}{n}\right)^k - \left(1 - \frac{2}{n}\right)^k, \end{aligned}$$

$$P(X_i X_j = 1) = 1 - 2 \left(1 - \frac{1}{n}\right)^k + \left(1 - \frac{2}{n}\right)^k.$$

Thus

$$E[X_i] = 1 - \left(1 - \frac{1}{n}\right)^k, \quad \text{Var}(X_i) = \left(1 - \frac{1}{n}\right)^k \left(1 - \left(1 - \frac{1}{n}\right)^k\right),$$

and

$$\text{Cov}(X_i, X_j) = 1 - 2 \left(1 - \frac{1}{n}\right)^k + \left(1 - \frac{2}{n}\right)^k - \left(1 - \left(1 - \frac{1}{n}\right)^k\right)^2.$$

Consequently

$$E[X_1 + \dots + X_n] = n \left(1 - \left(1 - \frac{1}{n} \right)^k \right)$$

and

$$\begin{aligned} \text{Var}(X_1 + \dots + X_n) &= n \left(1 - \frac{1}{n} \right)^k \left(1 - \left(1 - \frac{1}{n} \right)^k \right) \\ &+ n(n-1) \left(1 - 2 \left(1 - \frac{1}{n} \right)^k + \left(1 - \frac{2}{n} \right)^k - \left(1 - \left(1 - \frac{1}{n} \right)^k \right)^2 \right). \end{aligned}$$

Example 12

10 couples are randomly seated at a round table. Let X be the number of couples that are seated together. Find $\text{Var}(X)$.

For $i = 1, \dots, 10$, let $X_i = 1$ if the i -th couple are seated together. Then $X = X_1 + \dots + X_{10}$. For $i = 1, \dots, 10$,

$$P(X_i = 1) = \frac{2(18)!}{(19)!} = \frac{2}{19}.$$

For $i \neq j$,

$$P(X_i = 1, X_j = 1) = \frac{2^2(17)!}{(19)!} = \frac{4}{19 \cdot 18}.$$

Example 12

10 couples are randomly seated at a round table. Let X be the number of couples that are seated together. Find $\text{Var}(X)$.

For $i = 1, \dots, 10$, let $X_i = 1$ if the i -th couple are seated together. Then $X = X_1 + \dots + X_{10}$. For $i = 1, \dots, 10$,

$$P(X_i = 1) = \frac{2(18)!}{(19)!} = \frac{2}{19}.$$

For $i \neq j$,

$$P(X_i = 1, X_j = 1) = \frac{2^2(17)!}{(19)!} = \frac{4}{19 \cdot 18}.$$

Thus for $i = 1, \dots, 10$,

$$E[X_i] = \frac{2}{19}, \quad \text{Var}(X_i) = \frac{2}{19} \frac{17}{19}.$$

For $i \neq j$,

$$\text{Cov}(X_i, X_j) = \frac{4}{19 \cdot 18} - \frac{4}{19^2}.$$

$$\begin{aligned} \text{Var}(X) &= \text{Var}\left(\sum_{i=1}^{10} X_i\right) \\ &= \sum_{i=1}^{10} \text{Var}(X_i) + \sum_{i \neq j} \text{Cov}(X_i, X_j) \\ &= \frac{20}{19} \frac{17}{19} + 90 \left(\frac{4}{19 \cdot 18} - \frac{4}{19^2} \right). \end{aligned}$$

Thus for $i = 1, \dots, 10$,

$$E[X_i] = \frac{2}{19}, \quad \text{Var}(X_i) = \frac{2}{19} \frac{17}{19}.$$

For $i \neq j$,

$$\text{Cov}(X_i, X_j) = \frac{4}{19 \cdot 18} - \frac{4}{19^2}.$$

$$\begin{aligned} \text{Var}(X) &= \text{Var}\left(\sum_{i=1}^{10} X_i\right) \\ &= \sum_{i=1}^{10} \text{Var}(X_i) + \sum_{i \neq j} \text{Cov}(X_i, X_j) \\ &= \frac{20}{19} \frac{17}{19} + 90 \left(\frac{4}{19 \cdot 18} - \frac{4}{19^2} \right). \end{aligned}$$