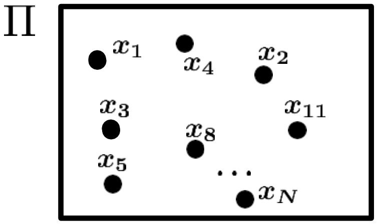
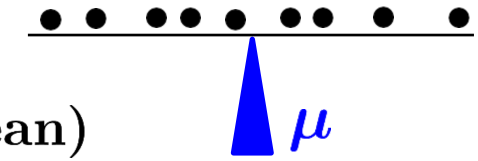
Finite Population Correction

(see problem 5.47 in page 184 of the textbook)

Let $\Pi = \{x_1, x_2, \ldots, x_N\}$ be a finite population of size N, and suppose we want to estimate an unknown population mean

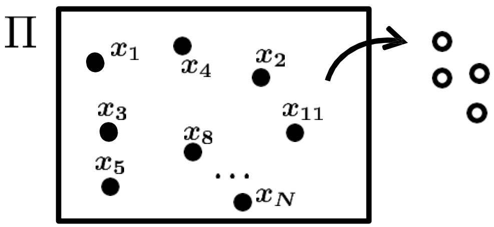


$$\boldsymbol{\mu = \frac{1}{N}\sum_{i=1}^{N} x_i}$$ **(Population mean)**

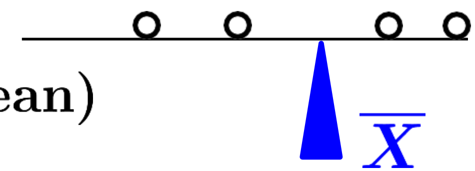


using a sample $(X_1, X_2, \ldots, X_n)$ of size n $(2 \leq n < N)$ drawn from Π.

Usually we use the sample mean



$$\boldsymbol{\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i}$$ **(Sample mean)**



as an unbiased estimator of μ based on the sample above.

$$\boldsymbol{E(\overline{X}) = \mu}$$ (Unbiasedness of the sample mean)

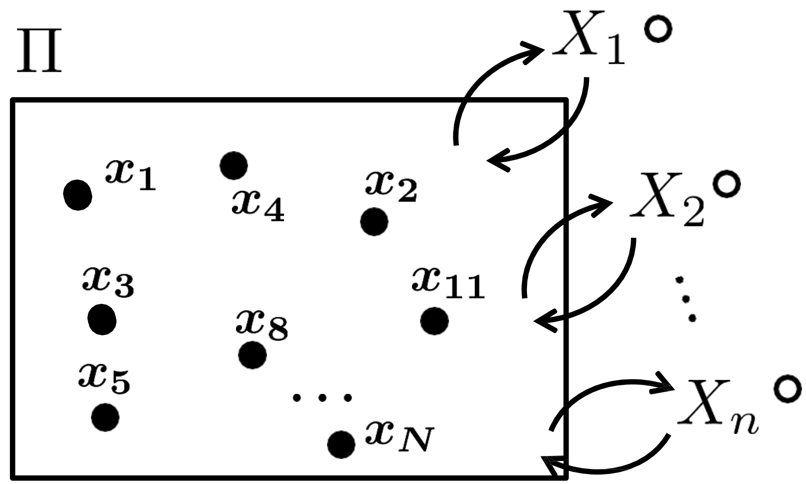
"$E(Y)$" denotes
"the expectation of Y".

We will check this later.

(The sample mean $\boldsymbol{\overline{X}}$ **is unbiased whether sampling is with or without replacement**.)

$$\boxed{\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2 \ \text{(Population variance)}}$$

Sampling with replacement



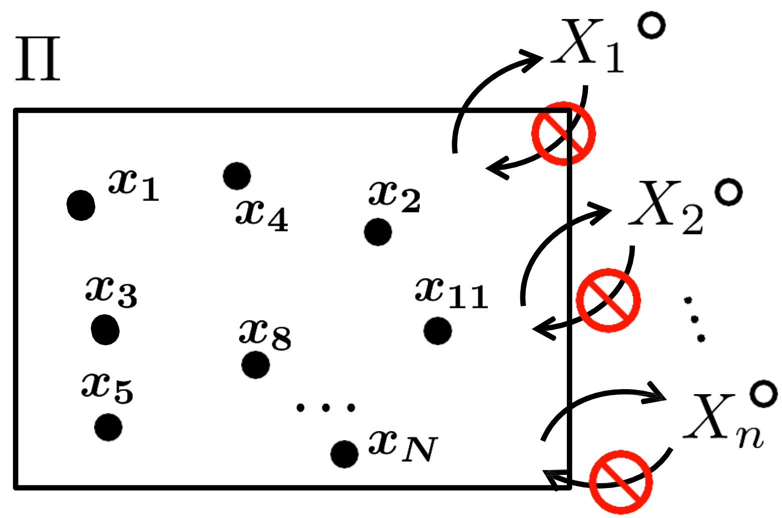
"$V(Y)$" denotes "the variance of Y".

$$V(\overline{X}) = \frac{\sigma^2}{n}$$

$$E(\overline{X}) = \mu$$

The expectation of $\overline{X}$ is the unknown population mean μ, i.e., $\overline{X}$ is unbiased, whether sampling is with or without replacement.

Sampling without replacement



$$E(\overline{X}) = \mu$$

$$V(\overline{X}) = \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right)$$

The constant $\boldsymbol{FPC = \sqrt{\dfrac{N-n}{N-1}}}$

is called *the finite population correction factor(FPC).*

We can use FPC to construct confidence intervals for means or proportions when sampling is without replacement from a finite population (see CHAPTER 6 Estimation Theory in the textbook).

For example, suppose the population distribution is normal with known standard deviation σ, $N = 500$ and $n = 100$. Then the 95% confidence interval for the unknown population mean μ is

$$\overline{X} \pm 1.96 \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}},$$

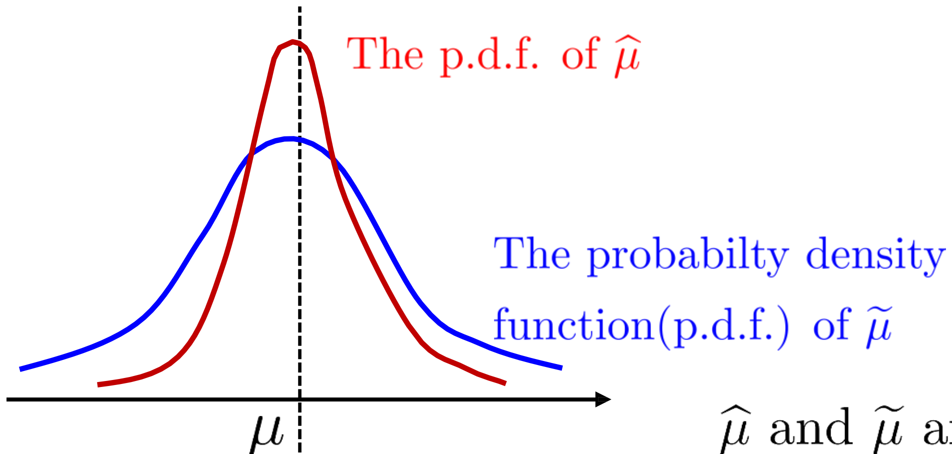
where $FPC = \sqrt{\dfrac{500-100}{500-1}} \doteqdot 0.8953$.

Population size:
$N = 10,000$

Sample size (n)	FPC
5	0.9998
10	0.9995
50	0.9975
100	0.9950
300	0.9849
500	0.9747
$1,000$	0.9487
$3,000$	0.8367
$5,000$	0.7071
$7,000$	0.5477
$9,000$	0.3162
$9,500$	0.2236

But why is $V(\overline{X})$ expressed as $\frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right)$ when sampling is without replacement from a finite population?

The accuracy of an unbiased estimator is usually measured by its variance.



$\widehat{\mu}$ and $\widetilde{\mu}$ are unbiased estimators of μ.

$$E(\widehat{\mu}) = E(\widetilde{\mu}) = \mu$$

$$V(\widehat{\mu}) < V(\widetilde{\mu})$$

$\Longrightarrow$ $\widehat{\mu}$ is more accurate than $\widetilde{\mu}$.

If sampling is **with replacement**, the variance of the sample mean $\overline{X}$ is given by

$$\boldsymbol{V(\overline{X}) = \frac{\sigma^2}{n}}$$

as is well known (but we will check this later).

If sampling is **without replacement**, on the other hand, the variance of $\overline{X}$ is given by

$$\boldsymbol{V(\overline{X}) = \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right)},$$

The constant $\sqrt{\dfrac{N-n}{N-1}}$ is named as *the finite population correction factor*.

We will check this later.

which is less than $\dfrac{\sigma^2}{n}$ since $\dfrac{N-n}{N-1} < 1$.

Recall the range of values for n: $\quad 2 \leq n < N$

This shows that we should adopt sampling without replacement when we take a sample from a finite population and then estimate μ by the sample mean. If, however, N is large and n/N is small, then the fraction

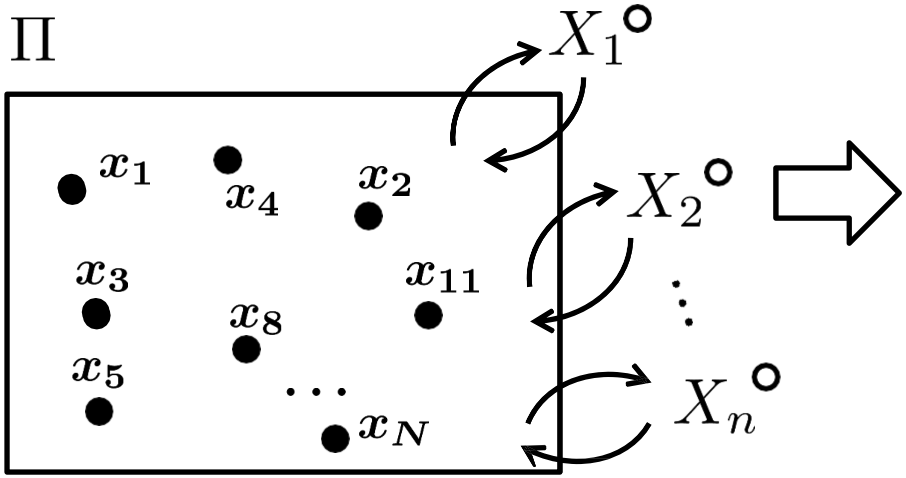
$$\frac{N-n}{N-1} = \frac{1-\frac{n}{N}}{1-\frac{1}{N}}$$

is nearly equal to 1, resulting in the negligible difference between sampling with and without replacement:

$$\frac{\sigma^2}{n} \fallingdotseq \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right)$$

The remainder of the slides is devoted to giving proofs of the following (1)–(4):

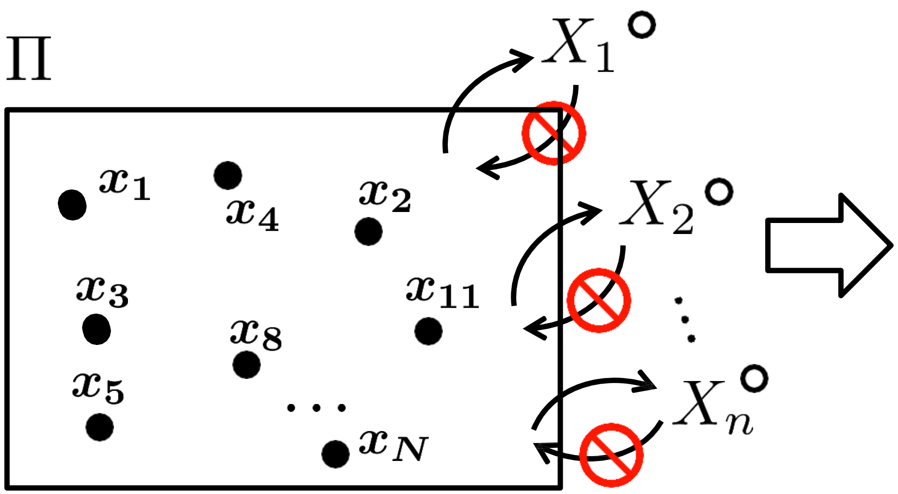
Sampling with replacement



$$\boldsymbol{E(\overline{X}) = \mu} \qquad \cdots (1)$$

$$\boldsymbol{V(\overline{X}) = \frac{\sigma^2}{n}} \qquad \cdots (2)$$

Sampling without replacement



$$\boldsymbol{E(\overline{X}) = \mu} \qquad \cdots (3)$$

$$\boldsymbol{V(\overline{X}) = \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right)}$$

$$\cdots (4)$$

The equalities (1)(2) and (3)(4) are derived on the basis of sampling with and without replacement, respectively.

The derivation of the former may be easy to understand, because you might have already learned about them in basic statistics courses. The derivation of the latter, on the other hand, would be slightly difficult to understand, because they require **some preparations** for their proofs.

(Sampling without replacement) (↓ The letter "P" denotes the probability of an event.)

$$\boldsymbol{E(\overline{X}) = \mu \cdots (3)}$$

$$\boxed{P(X_i = x_j) = \frac{1}{N} \quad \begin{pmatrix} i = 1, 2, \ldots, n \\ j = 1, 2, \ldots, N \end{pmatrix}}$$

↑ When sampling is without replacement, this is not obvious and should be verified.

$$\boxed{E(X_j) = \mu \quad (j = 1, 2, \ldots, n)}$$

(Sampling without replacement)

$$\boldsymbol{V(\overline{X})} = \frac{\boldsymbol{\sigma^2}}{\boldsymbol{n}}\left(\frac{\boldsymbol{N-n}}{\boldsymbol{N-1}}\right) \quad \cdots \textbf{(4)}$$

$$\boxed{V(X_j) = \sigma^2 \ (j = 1, 2, \ldots, n)}$$

Let $i, j = 1, 2, \ldots, n$ and $i \neq j$. Then

$$P(X_i = x_k, X_j = x_l) = \begin{cases} \dfrac{1}{N(N-1)} & (k \neq l) \\ 0 & (k = l) \end{cases}$$

(↑ A comma represents "and".)

for $k, l = 1, 2, \ldots, N$.

↑ Joint probabilities of X_i and X_j are different from those based on $X_1, X_2, \ldots, X_n$ from sampling with replacement.

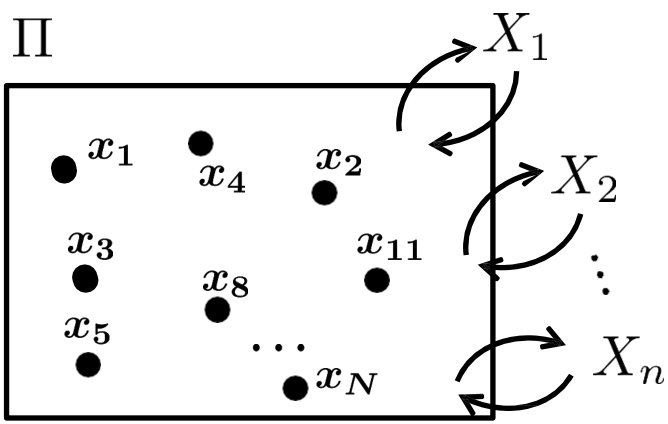
$$\boxed{Cov(X_i, X_j) = -\frac{\sigma^2}{N-1} \ (i \neq j)}$$

↑ The covariances of X_i and X_j $(i \neq j)$ are not zero since they are dependent, unlike in the case of sampling with replacement.

Now let's go through the proofs of (1)–(4) in order.

The proofs of (1) and (2).

$\Pi = \{x_1, x_2, \dots, x_N\}$: A finite population of size N.
$(X_1, X_2, \dots, X_n)$: A sample of size $n\,(< N)$ drawn from Π **with replacement**.



A random sampling procedure is one under which each selection of 1 out of N has the same probability $\frac{1}{N}$.

$$P(X_i = x_j) = \frac{1}{N} \quad \left(\begin{array}{l} i = 1, 2, \dots, n \\ j = 1, 2, \dots, N \end{array}\right)$$

The marginal probability distribution of X_i is the discrete uniform distribution.

$$P(X_i = x_k, X_j = x_l)$$
$$= P(X_i = x_k) P(X_j = x_l) = \frac{1}{N^2}$$
$(\uparrow \because X_1, X_2, \dots, X_n$ are independent.)

The joint probability of X_i and X_j is the product of their marginal probabilities.

$\boxed{\text{Theorem A1}}$ If $X_1, X_2, \ldots, X_n$ are drawn from $\Pi = \{x_1, x_2, \ldots, x_N\}$ **with replacement**, then

$$\boldsymbol{E(X_j) = \mu} \quad (j = 1, 2, \ldots, n)$$

$\boxed{\text{Proof}}$

$$E(X_j) = \sum_{i=1}^{N} x_i P(X_j = x_i)$$

($\uparrow$ by definition of the expectation of a discrete random variable)

$$= \sum_{i=1}^{N} x_i \cdot \frac{1}{N} = \frac{1}{N} \sum_{i=1}^{N} x_i$$

$$= \mu \quad (\leftarrow \text{by definition of the population mean}) \qquad \blacksquare$$

$\boxed{\text{Theorem A2}}$ If $X_1, X_2, \ldots, X_n$ are drawn from $\Pi = \{x_1, x_2, \ldots, x_N\}$ **with replacement**, then

$$\boldsymbol{E(\overline{X}) = \mu \quad \cdots (1)}$$

$\boxed{\text{Proof}}$

$$E(\overline{X}) \overset{(\downarrow \text{ by definition of the sample mean})}{=} E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \underset{(\uparrow \text{ by linearity of expectation})}{=} \frac{1}{n}\sum_{i=1}^{n} E(X_i)$$

$$= \frac{1}{n}\sum_{i=1}^{n} \mu \;=\; \frac{1}{n} \times n\mu \;=\; \mu$$

$(\uparrow\ E(X_i) = \mu\ (i = 1, 2, \ldots, n)$ by Theorem A1$)$ ∎

$\boxed{\text{Theorem A3}}$ If $X_1, X_2, \ldots, X_n$ are drawn from $\Pi = \{x_1, x_2, \ldots, x_N\}$ **with replacement**, then

$$\boldsymbol{V(X_j) = \sigma^2} \ \ (j = 1, 2, \ldots, n)$$

$\boxed{\text{Proof}}$

$$V(X_j) = E\left[\{X_j - E(X_j)\}^2\right] \text{(← by definition of the variance of a random variable)}$$

$$= E\left\{(X_j - \mu)^2\right\} \quad (\leftarrow E(X_j) = \mu \ (j = 1, 2, \ldots, n) \text{ by Theorem A1})$$

$$= \sum_{i=1}^{N} (x_i - \mu)^2 \, P(X_j = x_i) \quad \text{(← by definition of the variance of a discrete random variable)}$$

$$= \sum_{i=1}^{N} (x_i - \mu)^2 \cdot \frac{1}{N} \quad \left(\leftarrow \text{Recall the assumption } P(X_j = x_i) = \frac{1}{N} \ (i = 1, 2, \ldots, N)\right)$$

$$= \frac{1}{N} \sum_{i=1}^{N} (x_i - \mu)^2 = \sigma^2$$

(↑ by definition of the population variance) ∎

$\boxed{\text{Theorem A4}}$ If $X_1, X_2, \ldots, X_n$ are drawn from $\Pi = \{x_1, x_2, \ldots, x_N\}$ **with replacement**, then

$$\boldsymbol{V(\overline{X}) = \frac{\sigma^2}{n} \quad \cdots (2)}$$

$\boxed{\text{Proof}}$

$$V(\overline{X}) = V\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2} V\left(\sum_{i=1}^{n} X_i\right)$$

($\leftarrow$ The equality $V(cY) = c^2 V(Y)$ holds for any constant c and random variable Y (see (17) in p.78) .)

$$= \frac{1}{n^2}\left\{\sum_{i=1}^{n} V(X_i) + \sum_{i \neq j} Cov(X_i, X_j)\right\}$$

($\uparrow$ The equality $V(X+Y) = V(X)+V(Y)+2Cov(X,Y)$ holds for any random variables X and Y (see (51) in p.81). This equality can be generalized as above.)

($\uparrow$ Since the sampling is with replacement, X_i and X_j ($i \neq j$) are independent; this gives $E(X_i X_j) = E(X_i)E(X_j)$ and hence $Cov(X_i, X_j) = E(X_i X_j) - E(X_i)E(X_j) = E(X_i)E(X_j) - E(X_i)E(X_j) = 0$ (see (50) in p.81).)

$\boxed{\text{Theorem A4}}$ If $X_1, X_2, \ldots, X_n$ are drawn from $\Pi = \{x_1, x_2, \ldots, x_N\}$ **with replacement**, then

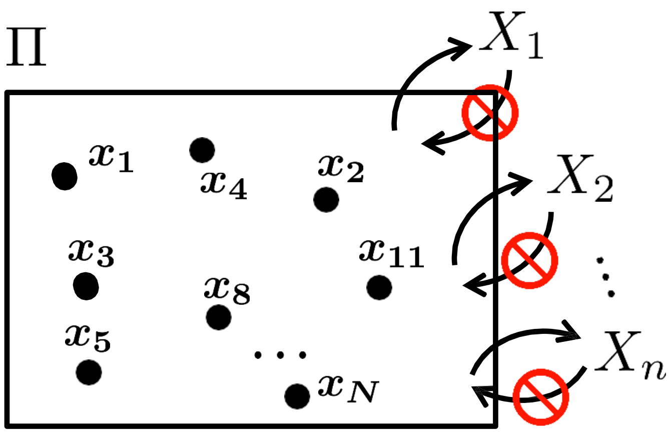
$$\boldsymbol{V(\overline{X}) = \frac{\sigma^2}{n} \quad \cdots (2)}$$

$\boxed{\text{Proof}}$ *(continued from the previous page.)*

$$\begin{aligned}
V(\overline{X}) &= \frac{1}{n^2} \sum_{i=1}^{n} V(X_i) \quad (\leftarrow V(X_i) = \sigma^2 \ (i = 1, 2, \ldots, n) \text{ by Theorem A3}) \\
&= \frac{1}{n^2} \sum_{i=1}^{n} \sigma^2 \\
&= \frac{1}{n^2} \times n\sigma^2 \\
&= \frac{\sigma^2}{n} \qquad \blacksquare
\end{aligned}$$

The proofs of (3) and (4).

$\Pi = \{x_1, x_2, \ldots, x_N\}$: A finite population of size N.
$(X_1, X_2, \ldots, X_n)$: A sample of size $n\,(< N)$ drawn from Π **without replacement**.



$$P(X_i = x_j) = \frac{1}{N} \quad \begin{pmatrix} i = 1, 2, \ldots, n \\ j = 1, 2, \ldots, N \end{pmatrix}$$

The marginal probability distribution of X_i is the discrete uniform distribution.

A random sampling procedure is one under which each selection of 1 out of $N - k$ $(k = 0, 1, 2, \ldots, N-1)$ has the same probability $\dfrac{1}{N-k}$.

Let $i, j = 1, 2, \ldots, n$ and $i \neq j$. Then

$$P(X_i = x_k, X_j = x_l) = \begin{cases} \dfrac{1}{N(N-1)} & (k \neq l) \\ 0 & (k = l) \end{cases}$$

for $k, l = 1, 2, \ldots, N$.

An example of

Let $i, j = 1, 2, \ldots, n$ and $i \neq j$. Then

$$P(X_i = x_k, X_j = x_l) = \begin{cases} \dfrac{1}{N(N-1)} & (\,k \neq l\,) \\ 0 & (\,k = l\,) \end{cases}$$

for $k, l = 1, 2, \ldots, N$.

:

$$\boldsymbol{\text{Why } P(X_1 = x_8, X_2 = x_1) = \frac{1}{N} \cdot \frac{1}{N-1} \text{ ?}}$$

$$\boldsymbol{P(X_1 = x_8, X_2 = x_1)} = \underline{P(X_1 = x_8)P(X_2 = x_1 \,|\, X_1 = x_8)}$$

($\uparrow$ $P(B|A)$ is the conditional probability of an event B given that an event A has occurred, i.e., $P(B|A) = \dfrac{P(B \cap A)}{P(A)}$ when $P(A) > 0$ (see (17) in p.7 of the textbook).)

$$= P(X_1 = x_8)\underline{P(X_2 \neq x_8, X_2 = x_1 \,|\, X_1 = x_8)}$$

($\uparrow$ $X_2 = x_1$ is chosen with uniform probability from the population Π without x_8.)

$$\boldsymbol{= \frac{1}{N} \cdot \frac{1}{N-1}}$$

An example of $\boxed{P(X_i = x_j) = \frac{1}{N} \begin{pmatrix} i = 1, 2, \ldots, n \\ j = 1, 2, \ldots, N \end{pmatrix}}$ **:** $\boxed{\textbf{Why } \boldsymbol{P(X_3 = x_5) = \frac{1}{N}} \textbf{ ?}}$

$$\boldsymbol{P(X_3 = x_5)} = P(X_1 \neq x_5, X_2 \neq x_5, X_3 = x_5)$$

X_1, X_2 and X_3 are drawn successively from the population Π without replacement, so the event $\{X_3 = x_5\}$ has the same meaning as the event $\{X_1 \neq x_5, X_2 \neq x_5, X_3 = x_5\}$.

The equality $P(A \cap B \cap C) = P(A)P(B|A)P(C|A \cap B)$ holds if $P(A \cap B) > 0$ (see (19) in p.7).

Suppose $X_1 = x_j$ $(j \neq 5)$ and x_j is excluded from Π; the population size changes to $(N-1)$. Since X_2 cannot be x_5, X_2 can take on $(N-2)$ values.

$$= P(X_1 \neq x_5)\overline{P(X_2 \neq x_5 | X_1 \neq x_5)}\underline{P(X_3 = x_5 | X_1 \neq x_5, X_2 \neq x_5)}$$

$$= \frac{\boldsymbol{N-1}}{\boldsymbol{N}} \cdot \frac{\boldsymbol{N-2}}{\boldsymbol{N-1}} \cdot \frac{\mathbf{1}}{\boldsymbol{N-2}}$$

Suppose $(X_1, X_2) = (x_j, x_k)$ $(j, k \neq 5; j \neq k)$ and x_j and x_k are excluded from Π; the population size changes to $(N-2)$.

$$= \frac{\mathbf{1}}{\boldsymbol{N}}$$

$\boxed{\text{Theorem B1}}$ If $X_1, X_2, \ldots, X_n$ are drawn from $\Pi = \{x_1, x_2, \ldots, x_N\}$ **without replacement**, then

$$\boldsymbol{E(X_j) = \mu} \quad (j = 1, 2, \ldots, n)$$

$\boxed{\text{Theorem B2}}$ If $X_1, X_2, \ldots, X_n$ are drawn from $\Pi = \{x_1, x_2, \ldots, x_N\}$ **without replacement**, then

$$\boldsymbol{E(\overline{X}) = \mu \quad \cdots (3)}$$

$\boxed{\text{Theorem B3}}$ If $X_1, X_2, \ldots, X_n$ are drawn from $\Pi = \{x_1, x_2, \ldots, x_N\}$ **without replacement**, then

$$\boldsymbol{V(X_j) = \sigma^2} \quad (j = 1, 2, \ldots, n)$$

We omit the proofs of the theorems above, because they are the same as for Theorems A1, A2 and A3, respectively.

$\boxed{\text{Theorem B4}}$ Let $i, j = 1, 2, \ldots, n$ and $i \neq j$. Then the covariance of X_i and X_j, defined by $Cov(X_i, X_j) = E\left[\{X_i - E(X_i)\}\{X_j - E(X_j)\}\right]$, is $\dfrac{-\sigma^2}{N-1}$, where $X_1, X_2, \ldots, X_n$ are drawn from $\Pi = \{x_1, x_2, \ldots, x_N\}$ **without replacement**.

$$\boldsymbol{Cov(X_i, X_j) = -\frac{\sigma^2}{N-1}\ (\, i \neq j\,)}$$

$\boxed{\text{Proof}}$ First, squaring both side of $\sum_{i=1}^{N}(x_i - \mu) = 0$, which is equivalent to $\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$, we get

$$\boldsymbol{\sum_{k \neq l}(x_k - \mu)(x_l - \mu) = -N\sigma^2}\,. \quad \cdots (A)$$

$$\left(\begin{array}{l} (\because) \quad 0 = \left\{\displaystyle\sum_{i=1}^{N}(x_i - \mu)\right\}^2 = \displaystyle\sum_{i=1}^{N}(x_i - \mu)^2 + \sum_{k \neq l}(x_k - \mu)(x_l - \mu)\,; \\ \displaystyle\sum_{k \neq l}(x_k - \mu)(x_l - \mu) = -\sum_{i=1}^{N}(x_i - \mu)^2 = -N \cdot \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2 = -N \cdot \sigma^2. \end{array}\right)$$

Next, by the definition of expectation we get

$\boxed{\text{Proof}}$ *(continued from the previous page.)*

$$\boldsymbol{Cov(X_i, X_j) = \frac{1}{N(N-1)} \sum_{k \neq l} (x_k - \mu)(x_l - \mu)} \qquad \cdots (B)$$

$$\boldsymbol{(\, i \neq j \,).}$$

$$(\because)\ Cov(X_i, X_j) = E\left\{(X_i - \mu)(X_j - \mu)\right\}$$

$$= \left(\sum_{k=l} + \sum_{k \neq l}\right) (x_k - \mu)(x_l - \mu) \underline{P(X_i = x_k, X_j = x_l)}$$

Recall the joint probabilities

$$P(X_i = x_k, X_j = x_l) = \begin{cases} \dfrac{1}{N(N-1)} & (\, k \neq l \,) \\ 0 & (\, k = l \,) \end{cases}$$

$$= \sum_{k \neq l} (x_k - \mu)(x_l - \mu) \frac{1}{N(N-1)} = \frac{1}{N(N-1)} \sum_{k \neq l} (x_k - \mu)(x_l - \mu)\,.$$

Substituting (A) into (B), we obtain the equality

$$Cov(X_i, X_j) = \frac{1}{N(N-1)} \times (-N\sigma^2) = -\frac{\sigma^2}{N-1}\,, \ \text{as required.} \quad \blacksquare$$

$\boxed{\text{Theorem B5}}$ (cf. Theorem 5-3 in p.156)

The variance of $\overline{X}$ is $\dfrac{\sigma^2}{n}\left(\dfrac{N-n}{N-1}\right)$, where $X_1, X_2, \ldots, X_n$ are drawn from $\Pi = \{x_1, x_2, \ldots, x_N\}$ **without replacement**.

$$\boldsymbol{V(\overline{X}) = \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right) \cdots (4)}$$

$\boxed{\text{Proof}}$

$$V(\overline{X}) = V\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2} V\left(\sum_{i=1}^{n} X_i\right)$$

The equality $V(cY) = c^2V(Y)$ holds for any constant c and random variable Y (see (17) in p.78).

$$= \frac{1}{n^2}\left\{\sum_{i=1}^{n} V(X_i) + \sum_{i \neq j} Cov(X_i, X_j)\right\}$$

The equality $V(X+Y) = V(X) + V(Y) + 2Cov(X, Y)$ holds for any random variables X and Y (see (51) in p.81). This equality can be generalized as above.

$\boxed{\text{Proof}}$ *(continued from the previous page.)*

$$V(\overline{X}) = \frac{1}{n^2}\left\{\sum_{i=1}^{n}\sigma^2 + \underline{\sum_{i\neq j}\left(-\frac{\sigma^2}{N-1}\right)}\right\} \qquad \text{(by Theorems B3 and B4)}$$

The operation $\sum\limits_{i\neq j}$ requires the exclusion of the case $(i,j)=(1,1),\ldots,(n,n)$; accordingly the constant $\left(-\frac{\sigma^2}{N-1}\right)$ must be added $n^2-n=n(n-1)$ times.

$$= \frac{1}{n^2}\left\{\underline{n\,\sigma^2 + n(n-1)\left(-\frac{\sigma^2}{N-1}\right)}\right\}$$

The term $n\sigma^2$ can be factored out.

$$= \frac{n\,\sigma^2}{n^2}\left\{1 + (n-1)\left(-\frac{1}{N-1}\right)\right\}$$

$$= \boldsymbol{\frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right)} \qquad \blacksquare$$