Поставим следующую задачу: произведено n взаимно независимых экспериментов со случайной величиной X. Получены числа X_1, X_2, … , X_n. По имеющимся данным определить математическое ожидание EX и дисперсию $DX=E((X-EX)^2)=E(X^2)-(EX)^2$.

Напомним в очередной раз, что это невозможно. И единственное, что удаётся сделать, это найти формулы определяющие значения EX и DX по числам X_1, X_2, … , X_n (формулы, определяющие параметры распределений по результатам экспериментов X_1, X_2, … , X_n называются «оценками») наилучшим образом с некоторой точки зрения. Самое простое решение заключается в том, что формулы берутся совпадающими с соответствующими им формулами для дискретной случайной величины, принимающей возможные значения X_1, X_2, … , X_n с равными вероятностями (равными $1/n$). Таблица распределения такой дискретной случайной величины имеет вид:

x_i:	X_1	X_2	…	X_n
p_i:	$1/n$	$1/n$	…	$1/n$

Тогда оценка математического ожидания, называемая «выборочное среднее» $\bar{x}=\frac{1}{n}\sum_{i=1}^{n}X_i$ является приближённым значением для истинного математического ожидания EX. Такое приближение является наилучшим с некоторых точек зрения. Во-первых, такая оценка параметра a является «несмещённой», то есть если вообразить, что X_1, X_2, … , X_n являются не числами, а взаимно независимыми случайными величинами, имеющими то же распределение, что и исследуемая случайная величина, то математическое ожидание этой оценки совпадёт с определяемым параметром. В рассматриваемой ситуации: $E(\bar{x})=E\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\frac{1}{n}E\left(\sum_{i=1}^{n}X_i\right)=\frac{1}{n}\sum_{i=1}^{n}EX_i=\frac{1}{n}\sum_{i=1}^{n}EX=\frac{1}{n}n\cdot EX=EX$, поскольку все EX_i=EX. Справедливости ради, заметим, что эта оценка не единственная из несмещённых оценок. Только что было записано, что математическое ожидание любой из случайных величин X_i тоже равно EX. В силу этого, другой несмещённой оценкой параметра a является, например, X_1. Во-вторых, оценка $\bar{x}=\frac{1}{n}\sum_{i=1}^{n}X_i$ является «состоятельной». Под этим термином предполагается, что при $n\to\infty$ оценка стремится по вероятности к определяемому параметру. В данном случае, это означает, что для любого $\varepsilon>0$ верно, что $\lim_{n\to\infty}P(|\bar{x}-EX|>\varepsilon)=0$. Заметим, что действительно $\lim_{n\to\infty}P\left(\left|\frac{1}{n}\sum_{i=1}^{n}X_i-EX\right|>\varepsilon\right)=0$ по теореме Чебышева (напомним, что для выполнения теоремы Чебышева требуется чтобы существовали дисперсии DX_i). Отметим также, что альтернативная оценка X_1 не удовлетворяет требованию состоятельности, поскольку выражение $|X_1-EX|$ не зависит от n и $\lim_{n\to\infty}P(|X_1-EX|>\varepsilon)=P(|X_1-EX|>\varepsilon)\neq 0$.

Аналогично для DX хочется предложить оценку $\frac{1}{n}\sum_{i=1}^{n}(X_i-EX)^2$, заменив в ней математическое ожидание EX доступным для вычислений средним значением $\bar{x}=\frac{1}{n}\sum_{i=1}^{n}X_i$. Тогда для дисперсии DX получится другая оценка, называемая «выборочная дисперсия» $s^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{x})^2=\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\frac{1}{n}\sum_{j=1}^{n}X_j\right)^2$. При этом величина $s=\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\frac{1}{n}\sum_{j=1}^{n}X_j\right)^2}$ называется «выборочное среднеквадратичное отклонение».

Проверим эту оценку на выполнение свойства несмещённости: найдём

$$E\left(s^2\right)=E\left(\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\frac{1}{n}\sum_{j=1}^{n}X_j\right)^2\right)=\frac{1}{n}\cdot E\left(\sum_{i=1}^{n}\left(X_i-\frac{1}{n}\sum_{j=1}^{n}X_j\right)^2\right)=$$

$$=\frac{1}{n}\cdot\sum_{i=1}^{n}E\left(\left(X_i-\frac{1}{n}\sum_{j=1}^{n}X_j\right)^2\right)=\frac{1}{n}\cdot\sum_{i=1}^{n}E\left((X_i)^2-X_i\frac{2}{n}\sum_{j=1}^{n}X_j+\left(\frac{1}{n}\sum_{j=1}^{n}X_j\right)^2\right)=$$

$$=\frac{1}{n}\cdot\sum_{i=1}^{n}\left(E\left((X_i)^2\right)-E\left(X_i\frac{2}{n}\sum_{j=1}^{n}X_j\right)+E\left(\frac{1}{n}\sum_{j=1}^{n}X_j\right)^2\right)=$$

Это сумма n одинаковых слагаемых, поскольку все X_i имеют одинаковые распределения, такие же, как X_1.

$$=\frac{1}{n}\cdot n\cdot\left(E\left(X_1^{\,2}\right)-E\left(X_1\frac{2}{n}\sum_{j=1}^{n}X_j\right)+E\left(\frac{1}{n}\sum_{j=1}^{n}X_j\right)^2\right)=E\left(X_1^{\,2}\right)-E\left(X_1\frac{2}{n}\sum_{j=1}^{n}X_j\right)+E\left(\frac{1}{n}\sum_{j=1}^{n}X_j\right)^2=$$

$$=E\left(X_1^{\,2}\right)-\frac{2}{n}\cdot E\left(X_1\sum_{j=1}^{n}X_j\right)+\frac{1}{n^2}\cdot E\left(\sum_{j=1}^{n}X_j\right)^2=$$

Теперь нужно понять следующее. Результаты разных испытаний независимые. А математическое ожидание независимых случайных величин равно произведению их математических ожиданий. Например, при $j>1$ верно $E(X_1X_j)=EX_1\cdot EX_j$. Поэтому для продолжения выкладок нужно отделить слагаемые, где есть произведения результатов испытаний с разными индексами от произведений результатов с одинаковыми индексами (квадратов результатов испытаний).

$$=E\left(X_1^{\,2}\right)-\frac{2}{n}\cdot E\left((X_1)^2+\sum_{j=2}^{n}\left(X_1\cdot X_j\right)\right)+\frac{1}{n^2}\cdot E\left(\sum_{j=1}^{n}X_j\right)^2=$$

Квадрат суммы можно получить раскрывая произведение обычным образом (умножая каждое слагаемое первой суммы на каждое слагаемое второй суммы).

$$=E\left(X_1^{\,2}\right)-\frac{2}{n}\cdot\left(E\left((X_1)^2\right)+E\left(\sum_{j=2}^{n}\left(X_1\cdot X_j\right)\right)\right)+\frac{1}{n^2}\cdot E\left(\sum_{j=1}^{n}\sum_{k=1}^{n}\left(X_j\cdot X_k\right)\right)=$$

$$=E\left(X_1^{\,2}\right)-\frac{2}{n}\cdot\left(E\left(X_1^{\,2}\right)+\sum_{j=2}^{n}E\left(X_1\cdot X_j\right)\right)+\frac{1}{n^2}\cdot\sum_{j=1}^{n}\sum_{k=1}^{n}E\left(X_j\cdot X_k\right)=$$

В двойной сумме необходимо отделить друг от друга слагаемые с одинаковыми индексами от слагаемых с разными индексами.

$$=E\left(X_1^{\,2}\right)-\frac{2}{n}\cdot E\left(X_1^{\,2}\right)-\frac{2}{n}\cdot\sum_{j=2}^{n}E\left(X_1\cdot X_j\right)+\frac{1}{n^2}\cdot\left(\sum_{j=1}^{n}E\left(X_j\cdot X_j\right)+\sum_{j=1}^{n}\sum_{k=1,k\neq j}^{n}E\left(X_j\cdot X_k\right)\right)=$$

$$=E\left(X_1^{\,2}\right)-\frac{2}{n}\cdot E\left(X_1^{\,2}\right)-\frac{2}{n}\cdot\sum_{j=2}^{n}\left(EX_1\cdot EX_j\right)+\frac{1}{n^2}\cdot\left(\sum_{j=1}^{n}E\left(X_j^{\,2}\right)+\sum_{j=1}^{n}\sum_{k=1,k\neq j}^{n}E\left(X_j\right)\cdot E\left(X_k\right)\right)=$$

В очередной раз обратим внимание, что в каждой из сумм все слагаемые одинаковые.

$$=E\left(X_1^{\,2}\right)-\frac{2}{n}\cdot E\left(X_1^{\,2}\right)-\frac{2}{n}\cdot\sum_{j=2}^{n}\left(EX_1\cdot EX_1\right)+\frac{1}{n^2}\cdot\left(\sum_{j=1}^{n}E\left(X_1^{\,2}\right)+\sum_{j=1}^{n}\sum_{k=1,k\neq j}^{n}E\left(X_1\right)\cdot E\left(X_1\right)\right)=$$

Нужно аккуратно подсчитать количества слагаемых в каждой из сумм.

$$=E\left(X_1^{\,2}\right)-\frac{2}{n}\cdot E\left(X_1^{\,2}\right)-\frac{2}{n}\cdot\sum_{j=2}^{n}\left(EX_1\cdot EX_1\right)+\frac{1}{n^2}\cdot\left(n\cdot E\left(X_1^{\,2}\right)+\left(n^2-n\right)E\left(X_1\right)\cdot E\left(X_1\right)\right)=$$

$$= E\left(X_1^2\right) - \frac{2}{n}\cdot E\left(X_1^2\right) - \frac{2}{n}\cdot(n-1)\cdot\left(EX_1\right)^2 + \frac{1}{n^2}\cdot\left(n\cdot E\left(X_1^2\right) + \left(n^2-n\right)\left(EX_1\right)^2\right) =$$

$$= E\left(X_1^2\right) - \frac{2}{n}\cdot E\left(X_1^2\right) - 2\cdot\left(EX_1\right)^2 + \frac{2}{n}\cdot\left(EX_1\right)^2 + \frac{1}{n}\cdot E\left(X_1^2\right) + \left(1-\frac{1}{n}\right)\left(EX_1\right)^2 =$$

$$= E\left(X_1^2\right) - \frac{2}{n}\cdot E\left(X_1^2\right) - 2\cdot\left(EX_1\right)^2 + \frac{2}{n}\cdot\left(EX_1\right)^2 + \frac{1}{n}\cdot E\left(X_1^2\right) + \left(EX_1\right)^2 - \frac{1}{n}\left(EX_1\right)^2 =$$

$$= E\left(X_1^2\right) - \frac{1}{n}\cdot E\left(X_1^2\right) - \left(EX_1\right)^2 + \frac{1}{n}\cdot\left(EX_1\right)^2 = E\left(X_1^2\right) - \left(EX_1\right)^2 + \frac{1}{n}\cdot\left(\left(EX_1\right)^2 - E\left(X_1^2\right)\right) =$$

$$= E\left(X_1^2\right) - \left(EX_1\right)^2 - \frac{1}{n}\cdot\left(E\left(X_1^2\right) - \left(EX_1\right)^2\right) = \left(1-\frac{1}{n}\right)\left(E\left(X_1^2\right) - \left(EX_1\right)^2\right) =$$

$= \left(1-\frac{1}{n}\right)DX = \frac{n-1}{n}DX \neq DX$. Оценка s^2 не прошла тест на несмещённость. В то же время ясно, как исправить эту оценку, чтобы её математическое ожидание было равно DX. Нужно создать «исправленную выборочную дисперсию»: $\left(s_1\right)^2 = \frac{n}{n-1}s^2 = \frac{n}{n-1}\cdot\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \frac{1}{n}\sum_{j=1}^{n}X_j\right)^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \frac{1}{n}\sum_{j=1}^{n}X_j\right)^2$. Такая оценка уже будет несмещённой: $E\left(s_1^2\right) = E\left(\frac{n}{n-1}s^2\right) = \frac{n}{n-1}E\left(s^2\right) = \frac{n}{n-1}\cdot\frac{n-1}{n}DX = DX$. Заметим, что оценка $\frac{1}{n}\sum_{i=1}^{n}\left(X_i - EX\right)^2$ была несмещённой, просто ей было никак не воспользоваться на практике (если конечно каким-то чудом не узнать истинное значение EX).

Соответственно исправляется и оценка среднеквадратического отклонения. Величина

$s_1 = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \frac{1}{n}\sum_{j=1}^{n}X_j\right)^2}$ называется «исправленное выборочное среднеквадратичное отклонение».