Студент: Елисей Евсеев
Группа: М 4141
Домашняя работа: №1
Дата: 7 марта 2022 г.

Задача 1. Найти ошибку в следующем рассуждении: "Пусть дана выборка $X_1...X_n$ из некоторого распределения. Поскольку при каждом элементарном исходе случайная величина $X_{[n]}$ совпадает с одним из элементов выборки, то $X_{[n]}$ имеет такое же распределение, как и X_1 обосновать ответ.

Решение: Ошибка кроется в следующем месте : "Поскольку при каждом элементарном исходе случайная величина $X_{[n]}$ совпадает с одним из элементов выборки..."Здесь $X_{[n]}$ - максимальная случайная величина в вариационном ряде, т.е. наборе из исходных одинаково распределенных случайных независимых величин нашей выборки, упорядоченных по неубыванию. И функция распределения k-ой порядковой статистики (в данном случае речь именно об n-ой) отличается от функции распределения первой порядковой статистики X_1.

Задача 2. Пусть F_n - эмпирическая функция распределения, построенная по выборке $X_1...X_n$ Найти:

а) $DF_n(y)$

б) $D(F_n(z) - F_n(y))$

Решение:

а) Как известно, эмпирическая функция распределения равна количеству X_i - случ. величин, чье значение меньше y.

$$F_n(y) = \frac{1}{n} \sum I\{X_i < y\} \tag{1}$$

На практике мы нашли матожидание этой величины и оно равняется функции распределения (общей):

$$E[F_n(y)] = F(y) \tag{2}$$

Теперь найдем дисперсию n независимых индикаторных случайных величин (пользуемся тем что величины независимы и дисперсия суммы равна сумме дисперсий и $D[cX] = c^2 D[X]$):

$$D[F_n(y)] = D[\frac{1}{n} \sum I\{X_i < y\}] = \frac{1}{n^2} \sum D[I\{X_i < y\}] = \frac{1}{n^2} n D[I\{X_i < y\}] \tag{3}$$

Всего индикаторных с.в. n штук поэтому вынесется n и найдем дисперсию отдельной индикаторной с.в. (пользуясь тем, что $I^2 = I$ - индикаторная случайная величина в квадрате):

$$D[I\{X_i < y\}] = E[(I\{X_i < y\})^2] - E^2[I\{X_i < y\}] = F(y) - F^2(y) \tag{4}$$

Окончательно соединим:

$$D[F_n(y)] = \frac{F(y) - F^2(y)}{n} = \frac{F(y)(1 - F(y))}{n} \tag{5}$$

б) Запишем дисперсию суммы с.в. , но учтем их зависимость:

$$D[(F_n(z) + (-F_n(y)))] = D[F_n(z)] + D[-F_n(y)] + 2cov(F_n(z), (-F_n(y))) \tag{6}$$

Раскроем формулу для ковариации через матожидание в удобной форме:

$$cov(X, Y) = E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y] \tag{7}$$

$$cov(F_n(z),(-F_n(y))) = E[F_n(z),(-F_n(y))] - E[F_n(z)]E[F_n(y)] \tag{8}$$

Так как внутри дисперсии это все раскрывается в сумму индикаторных величин, то зависимыми окажутся только n штук из них, где X_{ij} причем $i = j$ (так как там произведение внутри матожидания, то появится двойной индекс ij внутри первого члена ковариации). Далее внимательно посмотрим чему соотвествует отрезок $F_n(z)$ и $F_n(y)$ (отрезок на числовой прямой и нам надо выбрать пересечение этих двух величин), нам нужно рассмотреть случаи $z \geq y$ и $z < y$.

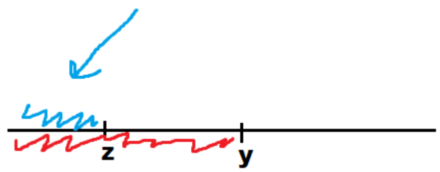

Рис. 1: $z < y$

$$E[F_n(z),(-F_n(y))] = \begin{cases} -F(z)F(y) & i \neq j \\ -F(z) & i = j, z < y \\ -F(y) & i = j, z \geq y \end{cases} \tag{9}$$

Тогда окончательно получим сгруппировав все вместе:

$$D[F_n(z) - F_n(y)] = \begin{cases} \frac{F_(z) - F(y) - (F^2(z) + F^2(y) - F(z)F(y) - F(y)F(z))}{n} & z \geq y \\ \frac{F_(y) - F(z) - (F^2(z) + F^2(y) - F(z)F(y) - F(y)F(z))}{n} & z < y \end{cases} \tag{10}$$

$$D[F_n(z) - F_n(y)] = \begin{cases} \frac{F_(z) - F(y) - (F(z) - F(y))^2}{n} & z \geq y \\ \frac{F_(y) - F(z) - (F(z) - F(y))^2}{n} & z < y \end{cases} \tag{11}$$

Задача 3. Пусть $X_1 ... X_n$ - выборка из распределения Бернулли с параметром p. Какое распределение имеет выборка $Y_1 ... Y_n$, где $Y_i = F(X_i)$, а $F(y)$ - функция распределения Бернулли.

Решение: Функция распределения Бернулли выглядит следующим образом:

$$F_\xi(x) = P(\xi < x) = \begin{cases} 0 & x \leq 0 \\ 1 - p & 0 < x \leq 1 \\ 1 & x > 1 \end{cases} \tag{12}$$

$$Y_i = \begin{cases} 0 & F(X_i) < 0 \\ 1 - p & 0 \leq F(X_i) < 1 \\ 1 & F(X_i) \geq 1 \end{cases} \tag{13}$$

Тогда в нашем случае для одинаково распределенных с.в. $Y_i = F(X_i)$, где F - распределение Бернулли, найдем распределение (подставив значения для F из (12) в точках 1 и 0):

Y_i	0	$1-p$	1
P	0	$1-p$	p