Решение семинара 1: Повторение теории вероятностей и математической статистики

Задача 11.

Доходность ценных бумаг на New York Фондовой бирже имеет нормальное распределение. В таблице приведены данные о доходности 10 видов ценных бумаг:

№	1	2	3	4	5	6	7	8	9	10	$\sum$
X	10	16	5	10	12	8	4	6	5	4	80
X^2	100	256	25	100	144	64	16	36	25	16	782

1) Найти точечные оценки для математического ожидания и дисперсии доходности.
2) Найти 90% доверительный интервал для математического ожидания доходности.

Решение:

1) $\bar{x}$ - выборочный аналог (оценка) математического ожидания для нашей случайной величины X

$$\hat{\mu} = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = 8,$$

Для оценки дисперсии используем:

$\hat{\sigma}^2 = s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = 15,78$ (эта оценка дисперсии обладает свойством несмещенности, т.е. $E\left(s^2\right) = \sigma^2$). Как показать несмещенность:

$$E\left(s^2\right) = E\left[\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \mu - (\bar{x} - \mu))^2\right] = \frac{1}{n-1}\left[E\left(\sum_{i=1}^{n}(x_i - \mu)^2\right) - 2E\left((\bar{x} - \mu)\sum_{i=1}^{n}(x_i - \mu)\right) + \sum_{i=1}^{n}E\left((\bar{x} - \mu)^2\right)\right] =$$

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

2) Доверительный интервал: $P(a < \mu < b) = 0,9$

Лемма Фишера: $\frac{(n-1)s^2}{\sigma^2} \sim \chi^2(n-1)$, $\frac{(\bar{x} - \mu)\sqrt{n}}{\sigma} \sim t_{(n-1)}$

$$P\left(-t_{(n-1)}^{\alpha/2} < \frac{\bar{x} - \mu}{s}\sqrt{n} < t_{(n-1)}^{\alpha/2}\right) = P\left(-t_{(n-1)}^{0.05} < \frac{\bar{x} - \mu}{s}\sqrt{n} < t_{(n-1)}^{0.05}\right)$$

В нашем случае нужно взять критические точки t-распределения с 9 степенями свободы, которые отсекают внутри себя 90% распределения (по таблицам: two-tailed t(9) для 10%). Т.к. распределение симметрично, этими точками будут:

$$P\left(-1,833 < \frac{\bar{x} - \mu}{s}\sqrt{n} < 1,833\right) = 0,9$$

$$P\left(-1,833 < \frac{8 - \mu}{\sqrt{15.78}}\sqrt{10} < 1,833\right) = 0,9$$

$$-1,833 < \frac{8 - \mu}{\sqrt{15.78}}\sqrt{10} < 1,833$$

Откуда искомый интервал составляет: $5,69 < \mu < 10,32$

Задача 12.

Пусть $X_1, ..., X_n$ - выборка из нормально распределенной генеральной совокупности, т.е. $X_i \sim N(\mu, \sigma^2), i = 1, ..., n$.

Построены следующие оценки для математического ожидания μ :

$$\mu_1 = \bar{X}, \ \mu_2 = X_1, \ \mu_3 = \frac{X_1}{2} + \frac{1}{2(n-1)}(X_2 + ... + X_n).$$

1) Какая из этих оценок является несмещенной?

2) Какая из этих оценок является наиболее эффективной?
3) Какая из этих оценок является состоятельной?

Решение:
1) *Несмещенность* означает, что математическое ожидание оценки $\hat{\mu}$ совпадает с оцениваемым теоретическим параметром μ.

$$E(\mu_1)=E(\bar{X})=E\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\frac{1}{n}\sum_{i=1}^{n}E(X_i)=\frac{1}{n}n\mu=\mu\ ,$$

$$E(\mu_2)=E(X_1)=\mu\ ,$$

$$E(\mu_3)=E\left(\frac{X_1}{2}+\frac{1}{2(n-1)}(X_2+...+X_n)\right)=\frac{\mu}{2}+\frac{1}{2(n-1)}(n-1)\mu=\mu$$

Таким образом, все три оценки являются несмещенными.
2) *Эффективность*: чем меньше дисперсия оценки, тем эта оценка считается эффективнее.

$$D(\mu_1)=D(\bar{X})=D\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\frac{1}{n^2}\sum_{i=1}^{n}D(X_i)=\frac{1}{n^2}n\sigma^2=\frac{\sigma^2}{n}\ ,$$

$$D(\mu_2)=D(X_1)=\sigma^2\ ,$$

$$D(\mu_3)=D\left(\frac{X_1}{2}+\frac{1}{2(n-1)}(X_2+...+X_n)\right)=\frac{\sigma^2}{4}+\frac{1}{4(n-1)^2}(n-1)\sigma^2=\frac{n\sigma^2}{4(n-1)}$$

$D(\mu_1)<D(\mu_3)<D(\mu_2)\Rightarrow$ первая оценка самая эффективная.
3) *Состоятельность* означает, что для больших выборок ($n\to\infty$) вероятность значимых отклонений величины оценки $\hat{\mu}$ от значения оцениваемого теоретического параметра μ равна нулю.

Достаточными условиями состоятельности оценки являются:

(1) $E\left(\hat{\mu}\right)-\mu\underset{n\to\infty}{\to}0$

(2) $D\left(\hat{\mu}\right)\underset{n\to\infty}{\to}0$

Достаточным условием состоятельности для всех трех оценок будет стремление дисперсии к нулю при увеличении выборки $D(\hat{\mu}_i)\xrightarrow[n\to\infty]{}0,\ i=1,2,3$ (первое условие, конечность математического ожидания, выполнено автоматически в силу несмещенности). Однако это имеет место только для первой оценки: $D(\mu_1)=\dfrac{\sigma^2}{n}\xrightarrow[n\to\infty]{}0$.