Составил: доц. Коротков Н.А.,
компьютерная вёрстка: доц. Коротков Н.А.

МАТЕМАТИЧЕСКАЯ СТАТИСТИКА (для Р7)

Памятка по ключевым вопросам теории для подготовки к экзамену

1. Определение выборки, или выборочной совокупности. Вариационный ряд. Объем выборки.

Выборка − это та часть генеральной совокупности, элементы которой подвергаются статистическому исследованию. ***Вариационный ряд*** − это последовательность всех элементов выборки, расположенных в порядке неубывания. ***Объем выборки***, т.е. число n её элементов, предполагается существенно меньшим, чем объём N генеральной совокупности.

2. Определение эмпирической функции распределения.

Эмпирической функцией распределения называется функция, определяемая формулой:

$$F_n(x) = \frac{\mu_n(x)}{n} = \frac{1}{n}\sum_{i=1}^{n} e(x - X_i),$$

здесь $\mu_n(x)$ − число элементов выборки, удовлетворяющих условию $X_i < x,$ а $e(x)$ − функция Хевисайда.

3. Свойства эмпирической функции распределения.

1. Условие ограниченности:

$$0 \le F(x) \le 1.$$

2. Условие неубывания: если $x_1 \le x_2,$ то

$$F(x_1) \le F(x_2).$$

3. $\lim\limits_{x\to-\infty} F(x) = 0, \quad \lim\limits_{x\to+\infty} F(x) = 1.$

4. Вероятность того, что значение, принятое случайной величиной X попадёт в промежуток $(a, b),$ вычисляется по формуле

$$P(a < X < b) = F(b) - F(a).$$

5. Функция распределения непрерывна слева, то есть

$$\lim_{x\to a^{-0}} F(x) = F(a).$$

4. Определение распределения χ^2.

Пусть $X_1,..., X_n$ – независимые случайные величины, каждая из которых имеет стандартное нормальное распределение. Тогда распределение случайной величины

$$Y = X_1^2 + \ldots + X_n^2$$

называется χ^2-распределением с n степенями свободы.

5. Определение выборочного начального момента $k-$го порядка. Определение выборочного среднего.

Выборочным начальным моментом $k-$го порядка называют величину

$$M_{nk} = M_{nk}(X) = \frac{1}{n}\sum_{i=1}^{n}(X_i)^k.$$

При $k = 1$ выборочный начальный момент называют выборочным средним и обозначают $\overline{X}$ Таким образом

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i .$$

6. Определение центрального выборочного момента $k-$го порядка. Определение выборочной дисперсии.

Выборочным центральным моментом $k-$го порядка называют величину

$$M_{nk} = M_{nk}(X) = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^k.$$

При $k = 2$ выборочный центральный момент называют выборочной дисперсией и обозначают $S^2 = S^2(X).$ Таким образом

$$S^2(X) = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^k.$$

7. Чему равны математическое ожидание и дисперсия выборочного среднего?

Математическое ожидание выборочного среднего совпадает с математическим ожиданием генеральной совокупности. Дисперсия выборочного среднего равна дисперсии генеральной совокупности, деленной на объем выборки.

8. Точечные и интервальные оценки параметров генеральной совокупности.

Точечная оценка − приближенное значение параметра генеральной совокупности, получаемое в результате обработки выборочной совокупности. Точечная оценка должна удовлетворять условиям ***несмещённости***, ***состоятельности*** и ***эффективности***.

Оценка называется ***несмещённой***, если её математическое ожидание совпадает со значением оцениваемого параметра.

Оценка называется ***состоятельной***, если она сходится по вероятности к оцениваемому параметру.

Несмещённая оценка называется ***эффективной***, если она имеет наименьшую дисперсию среди всех возможных несмещённых оценок параметра генеральной совокупности.

Оценка неизвестного параметра называется ***интервальной***, если она определяется концами интервала, внутри которого с заданной вероятностью находится оцениваемый параметр.

9. Определение γ-доверительного интервала.

Это случайный интервал, который содержит (накрывает) значение оцениваемого параметра с вероятностью, не меньшей γ.

10. Выписать доверительный интервал для математического ожидания в общей нормальной модели.

Пусть $\bar{X}$ – выборочное среднее, S^2 – выборочная дисперсия и $t_{\frac{1+\gamma}{2},n-1}$ – квантиль распределения Стьюдента с $(n-1)$ -степенью свободы, отвечающий вероятности $\frac{1+\gamma}{2}$. Доверительный интервал $(\bar{X} - \frac{S}{\sqrt{n-1}} t_{\frac{1+\gamma}{2},n-1}, \bar{X} + \frac{S}{\sqrt{n-1}} t_{\frac{1+\gamma}{2},n-1})$ содержит внутри математическое ожидание генеральной совокупности с вероятностью γ.

11. Сформулировать принцип, на котором основывается проверка статистической гипотезы.

Проверка статистической гипотезы основывается на принципе, в соответствии с которым маловероятные события считаются невозможными, а события, имеющие большую вероятность, считаются достоверными.

12. Определение ошибок первого и второго рода.

Ошибка, совершаемая при отклонении правильной основной гипотезы H_0 называется ошибкой первого рода.

Ошибка, совершаемая в том случае, когда принимается основная гипотеза, но в действительности верна альтернативная гипотеза, называется ошибкой второго рода.

13. Какова вероятность ошибки первого рода?

Вероятность ошибки первого рода равна вероятности попадания статистики критерия в критическую область, при условии, что верна гипотеза H_0.

14. Определение сходимости по вероятности.

Случайные величины $\eta_1, \eta_2, \ldots, \eta_n, \ldots$ называются ***сходящимися по вероятности*** к величине A (случайной или неслучайной), если для любого $\varepsilon > 0$ вероятность события $\{|\eta_n - A| < \varepsilon\}$ стремится к единице при $n \to \infty$, то есть:

$$\lim_{n\to\infty} P\{|\eta_n - A| < \varepsilon\} = 1.$$

Сходимость по вероятности символически записывают так:

$$\eta_n \xrightarrow[n\to\infty]{P} A.$$

15. Сформулировать теорему В.И. Гливенко.

Теорема Гливенко. Для любого значения аргумента эмпирическая (статистическая) функция распределения сходится по вероятности к теоретической функции распределения.

16. Сформулировать закон больших чисел и центральную предельную теорему.

Закон больших чисел: если случайные величины $\eta_1, \eta_2, \ldots, \eta_n, \ldots$ независимы, одинаково распределены и имеют математическое ожидание $M\eta_i = a,$ то

$$\frac{1}{n}(\eta_1 + \ldots + \eta_n) \xrightarrow[n\to\infty]{P} a$$

(где P означает сходимость по вероятности).

Центральная предельная теорема: если дополнительно к предыдущему существует $D\eta_i = \sigma^2 > 0,$ то при $n \to \infty$ распределение случайной величины

$$\frac{(\eta_1 + \eta_2 + \ldots + \eta_n - na)}{\sigma\sqrt{n}}$$

стремится к нормальному распределению $N(0,1)$.

17. Понятие случайных чисел.

Случайные числа – это числа, которые могут рассматриваться как значения независимых одинаково распределенных случайных величин. Как правило, имеются в виду значения случайных величин с равномерным распределением вероятностей в промежутке $[0,1]$.

18. Понятие псевдослучайных чисел

Псевдослучайные числа – это числа, получаемые по какому-либо алгоритму и, следовательно, не являющиеся случайными, но имитирующие их, так как свойства псевдослучайных чисел близки к свойствам случайных чисел.