Случайные величины

Случайные величины (с.в.) – численное значение, появляющееся в результате опыта, и принимающее произвольное значение из заранее определенного множества.

Существует два типа случайных величин: дискретные и непрерывные.

Дискретные случайные величины принимают в результате испытания одно из изолированного дискретного множества значений. Примеры дискретных случайных величин: оценка, полученная на экзамене, число попаданий в мишень в серии из 10 выстрелов и т. п.

Функция, связывающая значения дискретной случайной величины с их вероятностями, называется ее *распределением* (*законом распределения*). Обычно закон распределения записывается в виде таблицы вида

X	x_1	x_2	…	x_n	…	
P	p_1	p_2	…	p_n	…	

Пример Пусть X – число очков выпавшее на игральной кости при одном броске. Тогда, эта с.в. распределена по закону

X	1	2	3	4	5	6
P	1/6	1/6	1/6	1/6	1/6	1/6

Непрерывные случайные величины в результате испытания могут принимать любые значения из некоторого интервала.

Примеры непрерывных случайных величин: спортивный результат в беге или прыжках, рост и масса тела человека, сила мышц и др.

Функция распределения

Рассмотрим вероятность того, что случайная величина X окажется меньше или равной некоторому заданному числу x, т. е.

$$F(x) = P((X \leq x))$$

Свойства функции распределения совпадает со свойствами эмпирической функции распределения (см. 2.3.4)

1. $F(x)$ неубывающая функция.

2. $\lim\limits_{x \to -\infty} F(x) = 0$

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

Плотность распределения вероятностей

Для непрерывных случайных величин вводится понятие плотность распределения вероятностей, или "плотность вероятностей”, играющее исключительно важную роль при их описании.

Плотность вероятностей — это производная от функции распределения непрерывной случайной величины, т.е.

$$f(x) = \frac{dF(x)}{dx}$$

Числовые характеристики случайных величин

Распределение случайной величины, заданное в виде функции распределения или плотности вероятностей, полностью ее характеризует. Однако такая исчерпывающая характеристика случайной величины сложна и далеко не всегда необходима. Для решения многих практических задач не нужно знать распределение случайной величины, а достаточно иметь лишь некоторые обобщающие числовые характеристики этого распределения.

Пусть имеется дискретная случайная величина X с возможными значениями $x_1, x_2, \dots, x_k$ и вероятностями этих значений $p_1, p_2, \dots, p_k$.

1.**Выборочное среднее арифметическое**: $\bar{x}_{Вн} = \frac{1}{n}\sum\limits_{i=1}^{k} n_i x_i = \sum\limits_{i=1}^{k} x_i \frac{n_i}{n}$.

2.**Математическое ожидание** обозначает как M(X).

Математическое ожидание дискретной случайной величины равно сумме всех ее возможных значений, умноженных на вероятности этих значений:

$$M(X) = m_x = \sum_i x_i P(X = x_i) = \sum_i x_i p_i.$$

Для непрерывных случайных величин математическое ожидание определяется с помощью плотности вероятностей по формуле:

$$M(X) = m_x = \int_{-\infty}^{\infty} x f(x) dx.$$

Свойства математического ожидания

1. *M(C)=C,*

2. *M(CX)=CM(X).*

3. *M(X+Y)= M(X)+M(Y),*

4. *M(X×Y)= M(X)M(Y)*, если *X* и *Y* – независимые С.В.

3. **Дисперсией** случайной величины X называется математическое ожидание квадрата отклонений случайной величины от ее математического ожидания (сравните с определением п. 3.4.2). Дисперсия обозначается как *D(X)*, или σ^2.

$$D(X) = M\left((X - M(X))^2\right)$$

Для дискретных случайных величин

$$D(X) = \sum_i (x_i - M(X))^2 p_i,$$

Для непрерывных случайных величин

$$D(X) = \int_{-\infty}^{\infty} (x - M(X))^2 f(x) dx.$$

4. Положительный корень квадратный из дисперсии называется **средним квадратическим (стандартным) отклонением с**лучайной величины.

Эта величина обозначается, как

$\sigma x = \sqrt{D}$.

Свойства дисперсии

1. *D(C)=0,*

2. *D(CX)=* C^2*D(X).*

3. $D(X+Y)= D(X)+D(Y)$,если X и Y – независимые с.в.

Биномиальное распределение.

Пусть вероятность некоторого события **A** равна **P(A),** тогда вероятность события противоположного **q=1-P(**A).

Пусть испытание проводится **n** раз. Биномиальный закон позволяет рассчитать вероятность того, что среди **n** испытаний событие **A** произойдет **m** раз.

$$C_n^m = \frac{n(n-1)(n-2)...(n-m+1)}{m!}$$

$$P_{m,n} = C_n^m P^m(A)(1-P(A))^{n-m}$$

Задача: Лечение заболевания приводит к выздоровлению в 80%. Лечилось пятеро животных. Каковы вероятности того, что:

1. выздоровят все пятеро,
2. выздоровят четверо,
3. не выздоровит ни один,

Дано:

P(A)=0.8

n=5

m_1=5

m_2=4

m_3=0

$P_{5,5}$=? $P_{5,4}$=? $P_{5,0}$=?

Решение:

Применяют биномиальный закон распределения.

$$C_n^m = \frac{n(n-1)(n-2)...(n-m+1)}{m!}$$

$$P_{m,n} = C_n^m P^m(A)(1-P(A))^{n-m}$$

1. Рассчитывают число сочетаний $C_5^5 = \frac{5\cdot 4\cdot 3\cdot 2\cdot 1}{1\cdot 2\cdot 3\cdot 4\cdot 5} = 1$

Находят вероятность того, что выздоровят все пятеро животных:

$P_{5,5}=1\cdot0.8^5\cdot(1-0.8)^0=0.8^5=0.327$

2. Рассчитывают число сочетаний $C_5^4=\frac{5\cdot4\cdot3\cdot2}{1\cdot2\cdot3\cdot4}=5$

Находят вероятность того, что выздоровят четверо животных:

$P_{5,4}=5\cdot0.8^4\cdot(1-0.8)^1=5\cdot0.8^4\cdot0.2=0.409$

3. Рассчитывают число сочетаний $C_5^0=1$

Находят вероятность того, что не выздоровит ни одно животное:

$P_{5,0}=1\cdot0.8^0\cdot(1-0.8)^5=0.2^5=3.19\cdot10^{-4}$

Ответ: $P_{5,5}=0.327$; $P_{5,4}=0.409$; $P_{5,0}=3.19\cdot10^{-4}$

Распределение Пуассона.

Когда вероятность события очень мала ($P<0.1$) и исчисляется сотыми и тысячными долями единицы, то для описания такого рода распределений редких событий служит формула Пуассона.

$$P_{n,m}=\frac{\lambda^m}{m!}\ell^{-\lambda}=\frac{\lambda^m}{m!\,\ell^{\lambda}},$$

Закон Пуассона позволяет рассчитать вероятность того, что при **n** испытаниях нужное нам событие выпадает **m** раз.

Где: **λ=n p** -ожидаемое среднее значение;

m-частота ожидаемого события в **n** независимых испытаний;

e=2,7183 -основание натуральных логарифмов;

m!-факториал или произведение натуральных чисел **m!=** $1\cdot2\cdot3\cdot4...m$.

Задача: Предположим, что редкое заболевание встречается у 0.02% большой популяции. Из популяции производят случайную выборку в 10000 человек, которых проверяют на это заболевание. Каково ожидаемое число людей с заболеванием в этой выборке? Какова вероятность, что заболевание окажется у трёх человек?

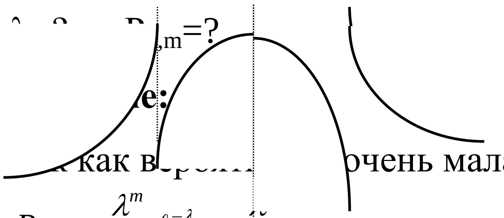
Дано:

$P=0.0002$

$n=10000$

$m=3$

$\lambda=2$ $P_{n,m}$=?



Решение:

Так как вероятность очень мала (P<0.1), применяем закон Пуассона:

$$P_{n(m)}=\frac{\lambda^m}{m!}\ell^{-\lambda}=\frac{\lambda^m}{m!\,\ell^{\lambda}},$$

1. Рассчитаем ожидаемое количество больных в данной выборке: $\lambda=n\cdot P$

$\lambda=10000\cdot0.0002=2$

2. Найдём вероятность того, что в этой выборке окажется трое больных.

$P_{n,3}=\frac{2^3\cdot\ell^{-2}}{3!}=\frac{8\cdot2.7^{-2}}{1\cdot2\cdot3}=0.36$

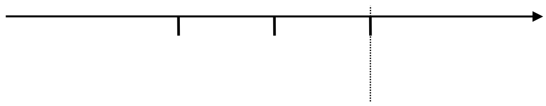
Ответ: $\lambda=2$; $P_{n,3}=0.36$

Нормальный закон распределения.

Для непрерывной случайной величины функция плотности вероятности рассчитывается по формуле: $f(x)=\frac{1}{\sigma\sqrt{2\pi}}\ell^{\frac{-(x_i-\bar{x})^2}{2\sigma^2}}$

График нормального распределения непрерывной случайной величины имеет вид:

f(x)



X_i

α a β

Вероятность того, что случайная величина лежит в интервале от α до β численно равна площади фигуры, заключенной между осью абсцисс и кривой, отвечающей нормальному закону. С помощью методов интегрального исчисления можно вычислить эту площадь. Площадь фигуры равна определенному интегралу от α до β от функции плотности вероятности. Тогда, вероятность того, что случайная величина лежит в интервале от α до β можно определить по формуле:

$$P(\alpha \le x \le \beta) = \int_{\alpha}^{\beta} \frac{1}{\sigma\sqrt{2\pi}} \ell^{\frac{-(x_i-\bar{x})^2}{2\sigma^2}} \mathrm{dx}$$

Вычисления упрощаются, если определенный интеграл от α до β от функции плотности вероятности представить как разность двух **F** функций, т. е.

$$P(\alpha \le x \le \beta) = \int_{\alpha}^{\beta} \frac{1}{\sigma\sqrt{2\pi}} \ell^{\frac{-(x_i-\bar{x})^2}{2\sigma^2}} \mathrm{dx} = F(\frac{\beta-\bar{x}}{\sigma}) - F(\frac{\alpha-\bar{x}}{\sigma})$$

Значения **F(t)**-функций определяются по таблице №1 (Значения интеграла вероятностей для разных значений t).

Задача:

Систолическое давление у женщин, страдающих гипертонической болезнью, имеет, согласно оценкам, среднее 158 мм. рт.ст. и стандартное отклонение 15 мм. рт.ст. В предположении, что систолическое давление является нормальной случайной величиной, оцените вероятность того, что давление находится между 141 и 177 мм.рт.ст. Какое количество женщин из 1000 имеет давление в этом интервале ?

Решение:

Дано:

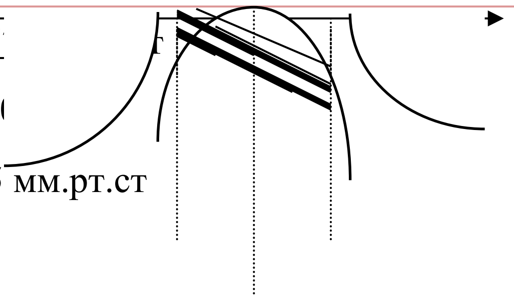
$\bar{X} = 158$ мм.рт.ст

α=141 мм.рт.ст

β=17

n=10

σ=15 мм.рт.ст



P(141≤ **X** ≤177)=?

Строят график нормального закона.

f(x)

*   *   *   *   *   * x_i

113   128   143   158   173   188   203

α β

1. Вероятность того, что случайная величина находится в интервале от 141 до 177мм.рт. ст. находят по формуле:

$$P(\alpha \le x \le \beta) = F(\frac{\beta - \overline{X}}{\sigma}) - F(\frac{\alpha - \overline{X}}{\sigma})$$

P(141≤ x ≤ 177)=F $(\frac{177-158}{15})$ – F$(\frac{141-158}{15}$)=F(1.27)+F(1.13)=0.3980+0.3708=0.7688

2. Чтобы найти, какое количество женщин имеет давление в этом интервале, используют формулу P=$\frac{m}{n}$, из которой находят m=n·P

$m=1000 \cdot 0.7688=768.8 \approx 769$

Ответ: $P=0.7688$; $m=769$