Теория Вероятности и Математическая статистика

Расчетная работа

Задание 2.1

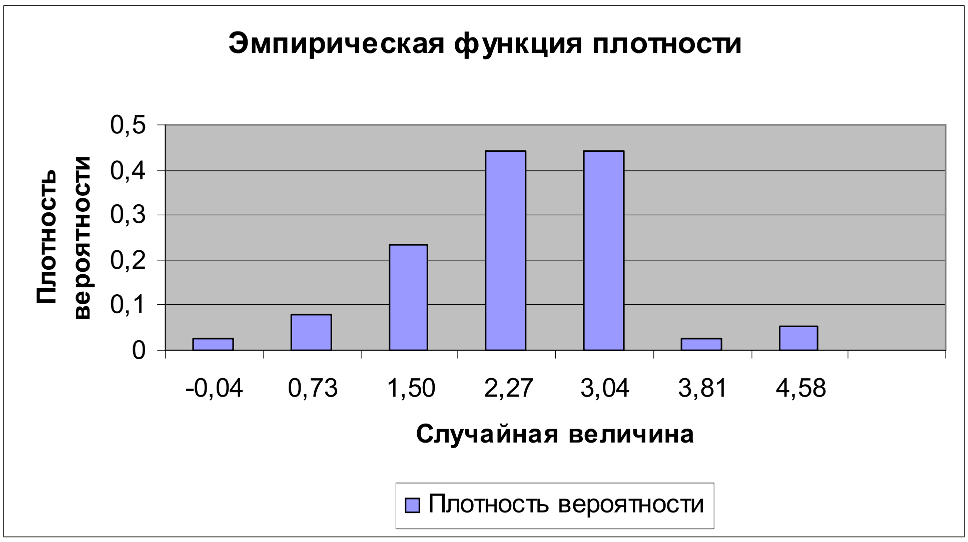
Построим группированный ряд наблюдений:

1. определим среди наблюдений в выборке минимальное Xmin=-0.04 и максимальное число Xmax=4.58
2. Разобьем диапазон [Xmin; Xmax] на m одинаковых интервалов, где $m=\log_2^n+1\approx6$. Отсюда ширина интервала Δ=(Xmax-Xmin)/m=0.771
3. Определим крайние точки каждого интервала (C_0, C_1, … , C_6). Причем за нижнюю величину интервала примем Xmin. Тогда $C_1= C_0+\Delta$, $C_2= C_0+2\Delta$, $C_3= C_0+3\Delta$, …

Таб.1

Интервал	Число элементов попавших в интервал	Частота	Эмпирическая функция плотности	Эмпирическая функция распределения	Теоретическая функция плотности
(-∞ -0,04]	1	0,02	0,025936058	0,02	0,026888372
(-0,04 0,73]	3	0,06	0,077808174	0,08	0,153110868
(0,73 1,50]	9	0,18	0,233424522	0,26	0,389641905
(1,5 2,27]	17	0,34	0,440912985	0,6	0,443142532
(2,27 3,04]	17	0,34	0,440912985	0,94	0,225236799
(3,04 3,81]	1	0,02	0,025936058	0,96	0,051162693
(3,81 4,58]	2	0,04	0,051872116	1	0,005193804
(4,58 +∞)	0	0	0	1	0

4. Далее построим гистограмму по формуле: $\frac{n_i}{n\Delta}$, исходя из вида данной кривой, сделаем предположение о том, что случайная величина имеет нормальное



распределение.

Задание 2.2

Метод моментов:

Суть метода моментов заключается в приравнивании теоретических моментов к их соответствующим эмпирическим оценкам: матожидание к среднему по выборке, дисперсию – к эмпирической дисперсии:

$$M[X] = \overline{X}$$
$$D[X] = S_x$$

M[X] = <u>2.01</u>

D[X] = <u>0.8925</u>

Метод квантилей:

Суть метода – приравнивание теоретических квантилей к эмпирическим квантилям:

$$\begin{cases} X^*_{p1} = X_{p1}, \\ X^*_{p2} = X_{p2} \end{cases}$$

где Xp_1 - теоретический квантиль уровня p_1, X^*p_1 - эмпирический квантиль уровня p_1.

Рассмотрим квантили двух уровней – 0,6 и 0,7

$$X^*_{0.6} = \frac{X_{(30)} + X_{(31)}}{2} = \frac{2.25 + 2.29}{2} = 2.27$$

$$X^*_{0.7} = \frac{X_{(35)} + X_{(36)}}{2} = \frac{2.49 + 2.61}{2} = 2.55$$

$$\begin{cases} \Phi(\frac{X^*_{0.6} - a}{\sigma}) = 0.6 \\ \Phi(\frac{X^*_{0.7} - a}{\sigma}) = 0.7 \end{cases} \Rightarrow \begin{cases} \frac{X^*_{0.6} - a}{\sigma} = 0.26 \\ \frac{X^*_{0.7} - a}{\sigma} = 0.53 \end{cases}$$

Отсюда $\hat{a}$ =2 , $\hat{\sigma}$ =1,0384

Метод максимального правдоподобия:

$$\varphi(x, a, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{\frac{-(x-a)^2}{2\sigma^2}}$$

$$L(X_1, X_2, .., X_n; a, \sigma) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{\frac{-(X_i - a)^2}{2\sigma^2}} = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \cdot e^{-\frac{\sum_{i=1}^{n}(X_i - a)^2}{(2\sigma^2)^n}}$$

$$\ln\{L(X_1, X_2, .., X_n; a, \sigma)\} = -\frac{n}{2}(\ln\sigma^2 + \ln(2\pi)) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i - \overline{X})$$

$$\frac{\partial \ln\{L\}}{\partial a} = 2\frac{\sum_{i=1}^{n}(X_i - a)^2}{(2\sigma^2)^n} = 0 \qquad \Rightarrow \qquad \hat{a} = \overline{X}$$

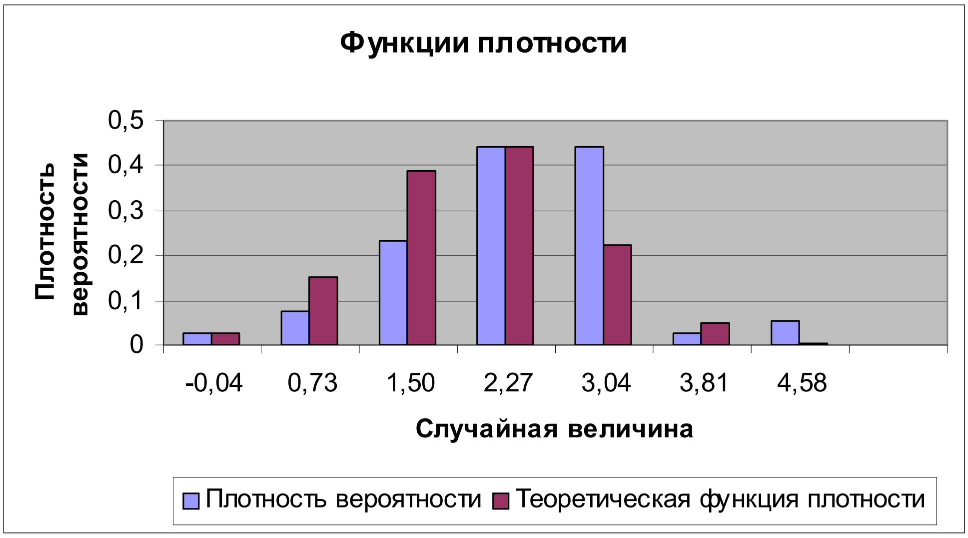
$$\frac{\partial \ln\{L\}}{\partial(\sigma^2)} = \frac{1}{2\sigma^4}\sum_{i=1}^{n}(X_i - a)^2 - \frac{n}{2\sigma^2} = 0 \qquad \Rightarrow \qquad \hat{\sigma}^2 = \frac{\sum_i (X_i - \overline{X})^2}{n}$$

M[X] =a= $\overline{X}$ =<u>2.01</u>

D[X] $= \sigma^2 = S_x{}^2 =$ <u>0.8925</u>

Задание 2.3

В качестве статистической оценки будем использовать оценки, полученные методом максимального правдоподобия(Данные указаны в Таб.1)



Задание 2.4

$$H_0: \quad X \sim N(a;\sigma^2) \;\Rightarrow\; F(x) = \Phi\left(\frac{x-a}{\sigma}\right)$$

$$H_1: \quad X \nsim N(a;\sigma^2)$$

1. Уровень значимости α=0.1
2. Сгруппируем интервалы с частотой появления элементов меньше 5, кроме 4 интервала (в противном случае нам бы не удалось бы вычислить Квыч.) В итоге получим следующие интервалы:

3.

Интервал	Число элементов	Частота	Теоретическая функция распределения	Вероятность
(-∞ 1,50]	13	0,26	0,276895845	0,276895845
(1,50 2,27]	17	0,34	0,619951808	0,343055963
(2,27 3,04]	17	0,34	0,885074259	0,265122451
(3,04 +∞)	3	0,06	0,998618694	0,113544434

4. Найдем вероятности p_i^0 с помощью оценок полученных ММП:

$$p_i^0 = P(X_i < X < X_{i+1}) = \Phi\left(\frac{X_{i+1} - \hat{a}}{\hat{\sigma}}\right) - \Phi\left(\frac{X_i - \hat{a}}{\hat{\sigma}}\right)$$

5. Найдем Квыч. по формуле: $\sum_{i=1}^{k}\left(\frac{\{n_i - n \cdot p_i^0\}^2}{n \cdot p_i^0}\right)$, где к – количество интервалов равно 4. В итоге: Квыч.=2.37

6. Найдем Крит. По формуле: $X_{1-\alpha}\left(\chi^2(k-r-1)\right) = X_{0.9}\left(\chi^2(1)\right) = \underline{2.71}$

Квыч.<Крит $\Rightarrow H_0$

Задание 2.5

1. $\hat{\sigma}^2 = S^2 x$

$$\overline{X} - \frac{t_\alpha Sx}{\sqrt{n-1}} < a < \overline{X} + \frac{t_\alpha Sx}{\sqrt{n-1}}$$

2. α=0.95

3. $t_{0.95} = X_{\frac{1+\alpha}{2}}[St(n-1)] = X_{0.95}[St(49)] = 2.009574$

4. В Итоге получим интервал для M[X]: 1.828<M[X]<2.3712

Задание 2.6

Возьмем интервал (0.73; 1.50]

Статистический метод: из Таб.1 возьмем соответствующую частоту - 0,18

Метод подстановки: суть метода заключается в подстановки оценки параметров в теоретическую функцию распределения, по которой мы в дальнейшем находим вероятность. Причем, в качестве оценки параметров будем использовать ММП. В итоге мы получим:

$$P(0.73 < X < 1.50) = \Phi\left(\frac{1.50 - \hat{a}}{\hat{\sigma}}\right) - \Phi\left(\frac{0.73 - \hat{a}}{\hat{\sigma}}\right) = \Phi\left(\frac{1.50 - 2.01}{0.8925}\right) - \Phi\left(\frac{0.73 - 2.01}{0.8925}\right) = 0{,}2087$$

Построим 80-% доверительный интервал:

$$w - U_\alpha\sqrt{\frac{w(1-w)}{n}} < P < w + U_\alpha\sqrt{\frac{w(1-w)}{n}}$$

$$U_\alpha: \quad \Phi(U_\alpha) = \frac{(1+\alpha)}{2}$$

$$U_{0.8}: \quad \Phi(U_{0.8}) = 0.9 \qquad \Rightarrow U_{0.8} = 1.28$$

В итоге получим интервал: 0,11<P<0,2495

Задание 2.7

$$\frac{n \cdot Sx}{X_{\frac{1+\alpha}{2}}\left[\chi^2(n-1)\right]} < \sigma^2 < \frac{n \cdot Sx}{X_{\frac{1-\alpha}{2}}\left[\chi^2(n-1)\right]}$$

1. α=0.8
2. $\begin{cases} X_{\frac{1-\alpha}{2}}\left[\chi^2(49)\right] = 36.82 \\ X_{\frac{1+\alpha}{2}}\left[\chi^2(49)\right] = 62.04 \end{cases}$

В итоге получим интервал: 0.7613< σ^2 <1,2828

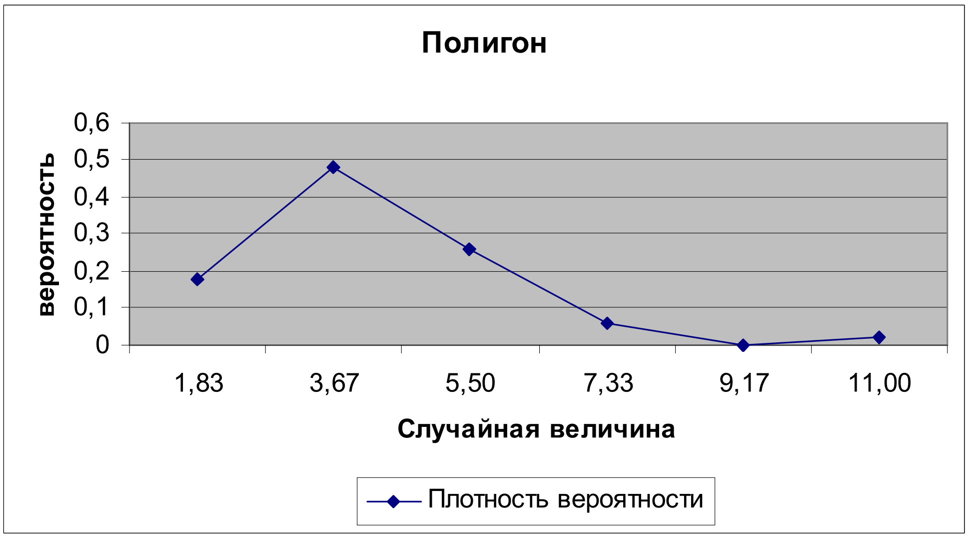
Задание 2.1

Построим группированный ряд наблюдений:

4. определим среди наблюдений в выборке минимальное Xmin =0 и максимальное число Xmax=11, дисперсия=3,69
5. Разобьем диапазон [Xmin; Xmax] на m одинаковых интервалов, где $m=\log_2^n+1\approx6$. Отсюда ширина интервала Δ=(Xmax-Xmin)/m=1,8334
6. Определим крайние точки каждого интервала (C_0, C_1, … , C_6). Причем за нижнюю величину интервала примем Xmin. Тогда $C_1= C_0+\Delta$, $C_2= C_0+2\Delta$, $C_3= C_0+3\Delta$, …

Интервал	Число элементов попавших в интервал	Частота	Эмпирическая функция распределения	Теоретическая функция вероятности
(0,00 1,83]	9	0,18	0,18	0,190364043
(1,83 3,67]	24	0,48	0,66	0,443427999
(3,67 5,50]	13	0,26	0,92	0,276120332
(5,50 7,33]	3	0,06	0,98	0,076834256
(7,33 9,17]	0	0	0,98	0,011978835
(9,17 11,00]	1	0,02	1	0,001188783
(11,00 +∞)	0	0	1	0

5. Далее построим полигон по формуле: $\frac{n_i}{n}$, исходя из вида данной кривой, сделаем предположение о том, что случайная величина имеет распределение Пуассона.



Задание 2.2

Метод моментов:

Суть метода моментов заключается в приравнивании теоретических моментов к их соответствующим эмпирическим оценкам: матожидание (лямбду) к среднему по выборке:

$M[X] = \lambda = \overline{X} = 3.06$

Метод квантилей:

Суть метода – приравнивание теоретических квантилей к эмпирическим квантилям:

$$X_{0.6} = X^*_{0.6}$$

$$X^*_{0.6} = \frac{X_{(30)} + X_{(31)}}{2} = 2$$

$$P(X \le X^*_{0.6}) = P(X \le 2) = P(X=0) + P(X=1) + P(X=2) = e^{-\lambda}(1 + \lambda + \frac{\lambda^2}{2}) = 0.6$$

Отсюда $\hat{\lambda} = 2.285$

Метод максимального правдоподобия:

$$P(X=m) = \frac{\lambda^m}{m!} \cdot e^{-\lambda}$$

$$L(X_1, X_2, .., X_n; \lambda) = \prod_{i=1}^{n} \frac{\lambda^{X_i}}{X_i!} \cdot e^{-\lambda} \Rightarrow$$

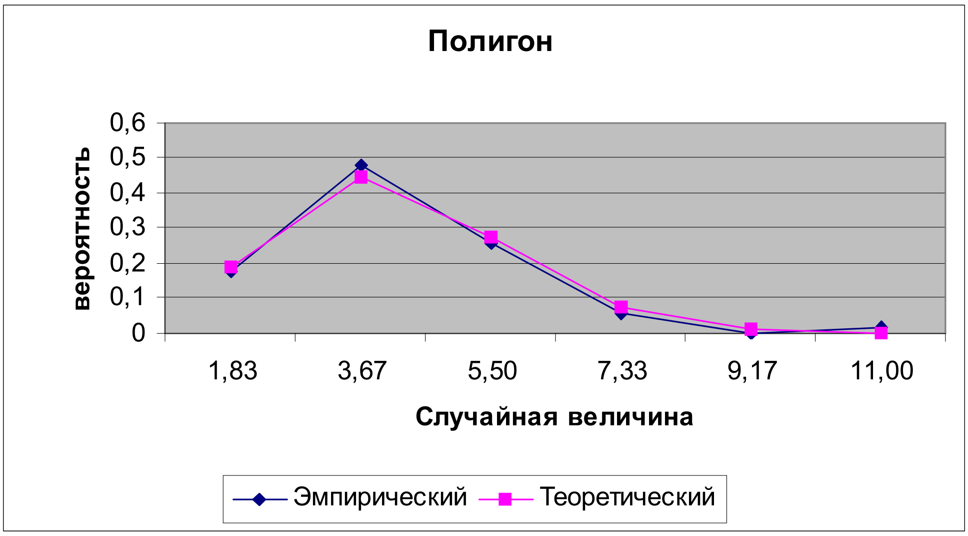
$$\ln\{L\} = \sum_i (X_i \ln\lambda - \ln(X_i!) - \lambda) = \ln\lambda \sum_i X_i - \sum_i \ln(X_i!) - n\lambda$$

$$\frac{\partial \ln L}{\partial \lambda} = \frac{1}{\lambda} \sum_i X_i - n = 0 \qquad \Rightarrow \hat{\lambda} = \overline{X}$$

$\hat{\lambda} = 3.06$

Задание 2.3

В качестве статистической оценки будем использовать оценки, полученные методом



максимального правдоподобия(Данные указаны в Таб.1)

Задание 2.4

$$H_0: \quad X \sim \Pi(\lambda) \qquad \Rightarrow P(x=m) = \frac{\lambda^m}{m!} \cdot e^{-\lambda}$$

$H_1: \quad X \nsim \Pi(\lambda)$

1. Уровень значимости α=0.1
2. Сгруппируем интервалы с частотой появления элементов меньше 5, кроме 4 интервала (в противном случае нам бы не удалось бы вычислить Квыч.) В итоге получим следующие интервалы:

Интервал	Число элементов	Частота	Вероятность
(0 1,83]	9	0,18	0,190364043
(1,83 3,67]	24	0,48	0,443427999
(3,67 5,50]	13	0,26	0,276120332
(5,50 +∞)	4	0,08	0,09014

3. Найдем вероятности p_i^0 с помощью оценок полученных ММП:

$$p_i^0 = P(X_i < X < X_{i+1}) = P([X_{i+1}]) + P([X_i]+1)$$

4. Найдем Квыч. по формуле: $\sum_{i=1}^{k}\left(\frac{\{n_i - n \cdot p_i^0\}^2}{n \cdot p_i^0}\right)$, где к – количество интервалов равно 4. В итоге: <u>Квыч.=0.3115</u>

5. Найдем Крит. По формуле: $X_{1-\alpha}\left(\chi^2(k-r-1)\right) = X_{0.9}\left(\chi^2(k-2)\right)$=4<u>.61</u>

Задание 2.5

$$\overline{X} - \frac{U_\alpha Sx}{\sqrt{n}} < M[X] < \overline{X} + \frac{U_\alpha Sx}{\sqrt{n}}$$

$U_{0.95} = 1.96$

<u>2,52<M[X]<3,592</u>

Задание 2.6

Возьмем интервал (3.67; 5.50]

<u>Статистический метод</u>: из Таб.1 возьмем соответствующую частоту - 0,26

<u>Метод подстановки</u>: суть метода заключается в подстановки оценки параметров в теоретическую функцию распределения, по которой мы в дальнейшем находим вероятность. Причем, в качестве оценки параметров будем использовать ОМП. В итоге мы получим:

P(3.67<X<5.50)=P(X=4)+P(X=5)=0,2761

Построим <u>80-% доверительный интервал</u>:

$$w - U_\alpha\sqrt{\frac{w(1-w)}{n}} < P < w + U_\alpha\sqrt{\frac{w(1-w)}{n}}$$

$$U_\alpha: \quad \Phi(U_\alpha) = \frac{(1+\alpha)}{2}$$

$$U_{0.8}: \quad \Phi(U_{0.8}) = 0.9 \qquad \Rightarrow U_{0.8} = 1.28$$

В итоге получим интервал: <u>0,185<P<0,339</u>

Задание 2.1

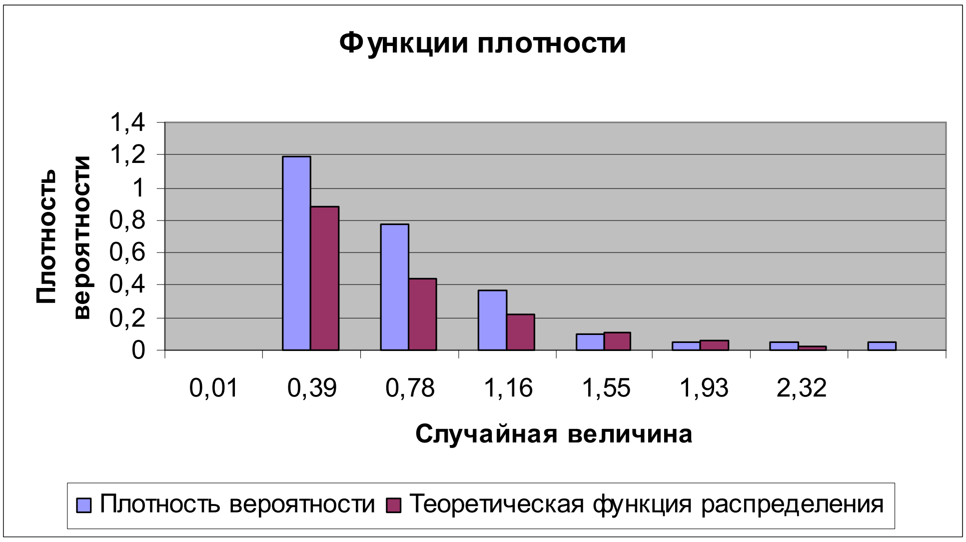
Построим группированный ряд наблюдений:

7. определим среди наблюдений в выборке минимальное Xmin=0.01 и максимальное число Xmax=2.32, дисперсия=0.28
8. Разобьем диапазон [Xmin; Xmax] на m одинаковых интервалов, где $m=\log_2^n+1\approx 6$. Отсюда ширина интервала Δ=(Xmax-Xmin)/m=0.384

Определим крайние точки каждого интервала (C_0, C_1, … , C_6). Причем за нижнюю величину интервала примем Xmin. Тогда $C_1= C_0+\Delta$, $C_2= C_0+2\Delta$, $C_3= C_0+3\Delta$, …

Интервал	Число элементов попавших в интервал	Частота	Эмпирическая функция плотности	Эмпирическая функция распределения	Теоретическая функция плотности
(0,01 0,39]	23	0,46	1,195432104	0,46	0,883805992
(0,39 0,78]	15	0,3	0,779629633	0,76	0,445107581
(0,78 1,16]	7	0,14	0,363827162	0,9	0,224167703
(1,16 1,55]	2	0,04	0,103950618	0,94	0,112896659
(1,55 1,93]	1	0,02	0,051975309	0,96	0,056857681
(1,93 2,32]	1	0,02	0,051975309	0,98	0,028635
(2,32 ∞)	1	0,02	0,051975309	1	0

6. Далее построим гистограмму по формуле: $\frac{n_i}{n\Delta}$, исходя из вида данной кривой, сделаем предположение о том, что случайная величина имеет показательное распределение.



Задание 2.2

Метод моментов:

Суть метода моментов заключается в приравнивании теоретических моментов к их соответствующим эмпирическим оценкам: матожидание к среднему по выборке, дисперсию – к эмпирической дисперсии:

$$M[X] = 1/\lambda = \overline{X} \quad \Rightarrow \quad \hat{\lambda} = \frac{1}{\overline{\overline{X}}} = 1.78$$

Метод квантилей:

Суть метода – приравнивание теоретических квантилей к эмпирическим квантилям:

$$\int_0^{X_p^*} \lambda e^{-\lambda x} dx = 1 - e^{-\lambda X_p^*} = p$$

Возьмем $X_{0.6}^* = \frac{X_{(30)} + X_{(31)}}{2} = \frac{0.64 + 0.61}{2} = 0.625 \quad \Rightarrow \quad \hat{\lambda} = 1.466$

Метод максимального правдоподобия:

$$L(X_1, X_2, .., X_n; \lambda) = \prod_{i=1}^{n} \lambda \cdot e^{-\lambda x} \Rightarrow \ln\{L\} = n \ln\lambda - \lambda X_i$$

$$\frac{\partial \ln L}{\partial \lambda} = \frac{n}{\lambda} - \sum_i X_i = 0 \qquad \Rightarrow \hat{\lambda} = \frac{1}{\overline{\overline{X}}} = 1.78$$

Задание 2.3

В качестве статистической оценки будем использовать оценки, полученные методом максимального правдоподобия(Данные указаны в Таб.)

Задание 2.4

$$H_0: \quad X \sim E(\lambda) \qquad \Rightarrow f(x, \lambda) = \lambda e^{-\lambda x}, x > 0$$

$$H_1: \quad X \nsim E(\lambda)$$

1. Уровень значимости α=0.1
2. Сгруппируем интервалы с частотой появления элементов меньше 5, кроме 4 интервала (в противном случае нам бы не удалось бы вычислить Квыч.) В итоге получим следующие интервалы:
3. Найдем вероятности p_i^0 с помощью оценок полученных ММП:

$$p_i^0 = P(X_i < X < X_{i+1}) = \int_{X_i}^{X_{i+1}} \hat{\lambda} e^{-\hat{\lambda} x} dx = e^{-\hat{\lambda} X_0} - e^{-\hat{\lambda} X_1}$$

Интервал	Число элементов	Частота	Вероятность
(0,00 0,39]	23	0,46	0,501022

(0,39 0,78]	15	0,3	0,249999
(0,78 1,16]	7	0,14	0,122509
(1,16 +∞)	5	0,1	0,126591

4. Найдем Квыч. по формуле: $\sum_{i=1}^{k}\left(\frac{\{n_i - n\cdot p_i^0\}^2}{n\cdot p_i^0}\right)$, где к – количество интервалов равно

4. В итоге: Квыч.=1.082

5. Найдем Крит. По формуле: $X_{1-\alpha}\left(\chi^2(k-r-1)\right)=X_{0.9}\left(\chi^2(k-2)\right)=4.\underline{61}$

Задание 2.5

$$\overline{X}-\frac{U_\alpha Sx}{\sqrt{n}}<M[X]<\overline{X}+\frac{U_\alpha Sx}{\sqrt{n}}$$

$U_{0.95}=1.96$

Интервал=0,4136-0,708

Задание 2.6

Возьмем интервал (0.39; 0.78]

Статистический метод: из Таб.1 возьмем соответствующую частоту - 0,3

Метод подстановки: суть метода заключается в подстановки оценки параметров в теоретическую функцию распределения, по которой мы в дальнейшем находим вероятность. Причем, в качестве оценки параметров будем использовать ОМП. В итоге мы получим: $\int_{X_0}^{X_1}\hat{\lambda}e^{-\hat{\lambda}x}dx=e^{-\hat{\lambda}X_0}-e^{-\hat{\lambda}X_1}$ ==0,25

Построим 80-% доверительный интервал:

$$w-U_\alpha\sqrt{\frac{w(1-w)}{n}}<P<w+U_\alpha\sqrt{\frac{w(1-w)}{n}}$$

$$U_\alpha:\quad \Phi(U_\alpha)=\frac{(1+\alpha)}{2}$$

$$U_{0.8}:\quad \Phi(U_{0.8})=0.9 \qquad \Rightarrow U_{0.8}=1.28$$

0,217<P<0,383

Задание 3

1. H_0 – две независимые выборки статистически однородны

   H_1 – две независимые выборки статистически не однородны

2. Разделим набор данных на 20(n) и 30(m) элементов соответственно: X и Y
3. Подсчитаем число успехов: число выполнения неравенства: $\{X_i < Y_j\}$ и числа «полу успехов» $\{X_i = Y_j\}$
4. Найдем Квыч, если суммарное число успехов U=272

$$\frac{\left|U-\frac{m\cdot n}{2}\right|}{\sqrt{\frac{n\cdot m(n+m+1)}{12}}}=\frac{\left|272-\frac{600}{2}\right|}{\sqrt{\frac{600(20+30+1)}{12}}}0.55$$

5. Найдем Ккрит.:

$$\begin{cases} U_\alpha = U_{0.05} \\ \Phi(U_{0.05}) = 1-\frac{\alpha}{2} = 0.975 \end{cases} \Rightarrow U_{0.05} = 1.96$$

Квыч.<Ккрит. $\Rightarrow H_0$

X			
0,16	0,51	0,35	1,15
0,26	0,16	0,31	0,28
0,47	1,10	0,05	2,10
0,73	0,16	0,14	2,32
0,81	0,57	0,37	0,41

Y		
0,13	0,64	0,92
1,15	0,87	0,20
0,08	0,02	0,01
0,11	0,05	0,26
0,73	0,07	0,13
0,89	0,44	0,05
0,68	0,46	0,01
0,68	0,51	0,61
0,25	0,51	1,21
0,66	1,51	1,78