№15 дәріс сабағы

χ^2 критерийі және оны түрлі болжамды тексеруге қолдану. Сызықтық және сызықтық емес регрессиялардың параметрлерін ең кіші квадраттар әдісімен анықтау.

Анықтама. Статистикалық болжам деп кездейсоқ шаманың үлестірімінің түрі немесе үлестірім параметрлері туралы алдын-ала жасалатын болжамды айтамыз. Мысалы, статистикалық болжамдарға мынадай болжамдар мысал бола алады:

1) Пуассон заңы бойынша үлестірілген бас жинақ;
2) екі қалыпты жинақтың дисперсиялары өзара тең.

Ұсынылған болжаммен қатар оған қарама-қарсы болжам да қарастырылады. Егер ұсынылған болжам жоққа шығарылса, онда оған қарама-қарсы болжам орындалады.

Анықтама. Нөлдік болжам деп ұсынылған H_0 болжамын атайды.

Анықтама. Альтернативті болжам деп нөлдік болжамға қарама-қарсы болжамды айтамыз.

Мысалы, нөлдік болжам «қалыпты үлестірімнің математикалық күтімі 10-ға тең» деген болжам болса, онда оған альтернативті болжам «қалыпты үлестірімнің математикалық күтімі 10-ға тең емес» деген болжам. Қысқаша былай жазамыз: $H_0 : a = 10$, $H_1 : a \neq 10$.

Анықтама. Жай болжам деп бір ғана сөйлемнен тұратын болжамды айтамыз.

Анықтама. Күрделі болжам деп шектеулі немесе шектеусіз жай болжамдардан тұратын болжамды айтамыз.

Ұсынылған болжам дұрыс та, бұрыс та болуы да мүмкін, сондықтан да тексеру қажет. Егер тексеру статистикалық әдістермен жүргізілсе, онда оны статистикалық деп атайды. Болжамдарды тексеру қорытындысында екі жағдайда дұрыс емес шешім қабылдануы мүмкін.

Анықтама. Бірінші текті қате деп – дұрыс болжамның жоққа шығарылуын айтамыз.

Анықтама. Екінші текті қате деп – бұрыс болжамның қабылдануын айтамыз.

Бірінші текті қате жіберу ықтималдығын маңыздылық деңгейі деп атаймыз және α деп белгілейміз. α маңыздылық деңгейі көбінде 0,05 немесе 0,01 мәндерін қабылдайды. Мысалы, $\alpha = 0,05$ дегеніміз барлығы 100 жағдайдың бесеуінде бірінші текті қате жіберу қаупі бар дегенді білдіреді.

Анықтама. Статистикалық критерий деп болжамды тексеру үшін қолданылатын кездейсоқ шаманы айтамыз.

Егер ол қалыпты үлестірілген болса, ол шаманы U немесе Z арқылы белгілейміз; ал егер Фишер заңы бойынша үлестірілсе - F арқылы; Стьюдент заңы бойынша үлестірілсе - T арқылы; «хи квадрат» заңы бойынша үлестірілсе χ^2 арқылы белгілейміз және т.с.с. Егер нақты үлестірімнің түрі туралы айтылмаса, онда жалпылық мақсатында ол шаманы K арқылы белгілейміз.

Анықтама. Бақылау мәні $K_{бақ}$ деп – таңдама бойынша есептелінетін критерийдің мәнін айтамыз.

3.7.2 Кризистік облыс. Болжамды қабылдау облысы. Кризистік нүктелер.

Анықталған критерий таңдалынып алынғаннан кейін оның барлық мүмкін мәндер жиынын екі қиылыспайтын ішкі жиындарға бөлеміз: оның біреуінің құрамында нөлдік болжам жоққа шығарылатын критерийдің мәні бар, ал екіншісінде - нөлдік болжам қабылданатын критерийдің мәні бар.

Анықтама. Кризистік облыс деп - нөлдік болжам жоққа шығаратын критерийдің мәндерінің жиынтығын айтамыз.

Анықтама. Болжамды қабылдау облысы деп - нөлдік болжамды қабылдайтын критерийдің мәндерінің жиынтығын айтамыз.

Анықтама. Кризистік облысты болжамды қабылдау облысынан бөліп тұратын нүктелерді *кризистік нүктелер* деп атаймыз.

Кризистік облыстар біржақты (оң жақты және сол жақты) және екіжақты кризистік облыстар болып бөлінеді.

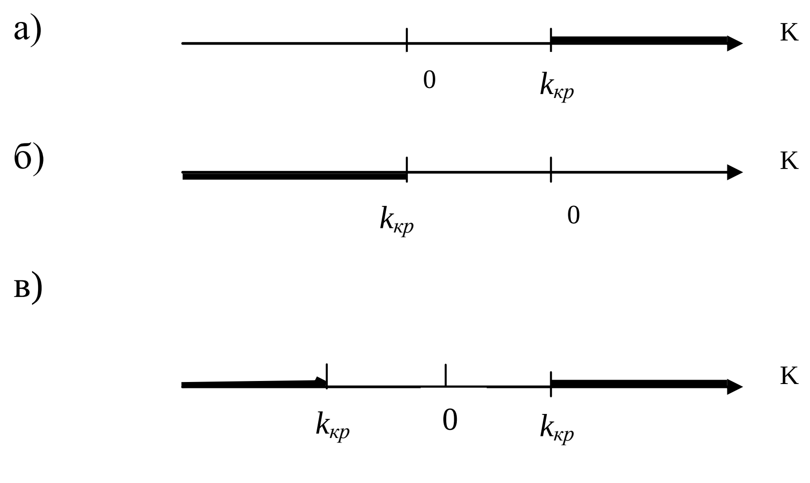
Анықтама. Оң жақты кризистік облыс деп $K > k_{кр}$ теңсіздігімен анықталатын облысты айтамыз, мұндағы $k_{кр}$ -оң сан (20а-сурет).

Анықтама. Сол жақты кризистік облыс деп $K < k_{кр}$ теңсіздігімен анықталатын облысты айтамыз, мұндағы $k_{кр}$ -теріс сан (20б-сурет).

Анықтама. Бір жақты кризистік облыс деп оң жақты және сол жақты кризистік облысты айтамыз.

Анықтама. Екі жақты кризистік облыс деп $K < k_1,\ K > k_2$, мұндағы $k_2 > k_1$, теңсіздіктерімен анықталатын облысты айтамыз.

Дербес жағдайда, егер кризистік нүкте нөлге қарағанда симметриялы болса, екі жақты кризистік облыс $|K| > k_{кр}$ теңсіздігімен анықталады(20в-сурет).



20-сурет

3.7.5 Статистикалық болжамдарды тексерудің логикалық схемасы

Шығарылатын есептердің өзіндік мінездемелері мен қойылатын талаптарына байланысты статистикалық критерийлер әртүрлі болып келеді. Бірақ оларды логикалық схеманың жалпылығы біріктіреді:

1) H_0 нөлдік болжамы мен H_1 альтернативтік болжамын ұсыну;

2) критерийдің α маңыздылық деңгейінің шамасы беріледі;

3) H_0 болжамын тексеру үшін критерийдің z статистикасы таңдалынып алынады;

4) H_0 болжамы ақиқат болатындай, z статистикасының таңдамалық үлестірімі анықталады;

5) альтернативті болжамға байланысты V_k кризистік облысы $z > z_{1-\alpha}$, $z < z_\alpha$ теңсіздіктерінің бірімен анықталады, немесе $z > z_{1-\alpha/2}$ және $z < z_{\alpha/2}$ теңсіздіктер жинағымен анықталады;

6) бақылаудың таңдамасын алады және критерийдің статистикасының z_b таңдамалық мәнін есептейді;

7) статистикалық шешім қабылданады:

егер $z_b \in V$ болса, онда бақылаудың нәтижелерімен келіспейтіндіктен H_0 болжамын жоққа шығарады;

егер $z_b \in V \setminus V_k$ болса, онда H_0 болжамын қабылдайды, яғни, H_0 болжамы бақылау нәтижелеріне қарама-қайшы келмейді деп есептеуге болады.

3.7.6 χ^2 келісімдік критерийі

Анықтама. Кездейсоқ шаманың заңының үлестірімі туралы болжамды тексеру үшін қолданылатын критерийлерді *келісімдік критерийлері* деп атаймыз.

Негізгі болжам H_0: X кездейсоқ шамасының үлестірім функциясы мынадай түрде: $F(x)$.

Сандық өсті мынадай аралықтарға бөлеміз:

$$(-\infty = a_0, a_1), [a_1, a_2], \ldots, [a_{r-1}, a_r = +\infty), \text{ мұндағы } a_1 < a_2 < \ldots < a_{r-1}.$$

H_0 болжамы: i-ші разрядқа $P_i = F(a_i) - F(a_{i1})$ ықтималдығы сәйкес келеді.

X кездейсоқ шамасының берілген $(x_1, x_2, \ldots, x_n)$ мәндерінің ішінен i-шы разрядқа m_i кездейсоқ саны сәйкес келеді $\left(\sum_{i=1}^{r} m_i = n\right)$.

Онда $\frac{m_i}{n}$ - таңдамалық мәннің i-ші интервалға түсу жиілігі. $\frac{m_i}{n}$ жиілігінің p_i санына жақындығы H_0 болжамының пайдасына шешіледі, ал едәуір айырмашылығы болса, онда H_0 болжамы жоққа шығарылады.

Кездейсоқ шама

$$\chi^2 = \sum_{i=1}^{2} \frac{(m_i - np_i)^2}{np_i},$$

H_0 болжамының сынақтың берілгендерімен сәйкес келетіндігін көрсетеді.

Критерийдің бақыланған $\chi^2 = \sum_{i=1}^{2} \frac{(m_i - np_i)^2}{np_i}$ формуласымен есептелінеді, ал кризистік жиын $S = (\chi^2_\alpha, +\infty)$ түрінде таңдап алынады, мұндағы χ^2_α мәні дайын кесте көмегімен табылады, $\gamma = r - 1 - t$ және маңыздылық деңгейі α белгілі болған жағдайларда, t дегеніміз - $F = F(x, \alpha_1, \alpha_2, \ldots, \alpha_t)$ функциясының параметрлерінің саны. Егер $\chi^2 > \chi^2_\alpha$ теңсіздігі ақиқат болса, онда H_0 болжамы α маңыздылық деңгейінде жоққа шығарылады. Кері жағдайда, H_0 болжамы сынақтың берілгендеріне қарама-қайшы болмайды.

3.8 Функционалдық, статистикалық және корреляциялық байланыстар

Анықтама. Шамалардың біреуінің өзгеруі екіншісінің үлестірімін өзгертетін болса, онда мұндай байланыс *статистикалық байланыс* деп аталады.

Анықтама. Егер шамалардың біреуінің өзгеруі екіншісінің орташа мәнін өзгертетін болса, онда бұл жағдайдағы статистикалық байланыс *корреляциялық байланыс* деп аталады.

Мысал 43. У дегеніміз – бидайдың түсімі болсын, ал Х - оған себілетін көңнің мөлшері. Аудандары бірдей жер бөліктеріне бірдей мөлшерде көң себілген, ал алынған түсімнің мөлшері әртүрлі, яғни, У Х-ке тәуелді функция емес. Бұл кездейсоқ факторлар (жауын-шашын, агротехника және т.б.) әсерінен деп түсіндіріледі. Жүргізілген тәжірибелер нәтижесінің көрсеткіші бойынша, орташа түсім себілетін көңнің мөлшеріне тәуелді. У пен Х-тің арасындағы байланыс корреляциялық байланыс.

Шартты орташалар

Әрбір X үшін бірнеше Y сәйкес келсін. Мысалы, X мәні: $x_1 = 2,$ ал Y мәні: $y_1 = 5,\ y_2 = 6,\ y_3 = 10$ болсын. Y шамасының қабылдануы мүмкін мәндерінің арифметикалық орташасы: $\bar{y}_2 = \frac{5+6+10}{3} = 7$ - шартты орташа.

Анықтама. $\bar{y}_x$ шартты орташа мәні деп $X=x$ болғандағы Y шамасының қабылдануы мүмкін мәндерінің арифметикалық орташасын айтамыз.

Анықтама. Y-тің X бойынша корреляциялық байланысы деп шартты орташалардың функционалдық байланысын айтамыз

$\bar{y}_x$ шамасының x бойынша: $$\bar{y}_x = f(x). \qquad (8)$$

(8) теңдеуі Y –тің X-ке байланысты регрессия теңдеуі деп атаймыз, $f(x)$ функциясын Y –тің X-ке байланысты регрессиясы деп атаймыз, ал оның графигін Y –тің X-ке байланысты регрессия сызығы деп атаймыз.

X-тің Y-ке корреляциялық байланысы дәл осылай анықталады.

3.9 Корреляция теориясының негізгі есептері

3.9.1 Регрессияның таңдамалық түзу сызық теңдеуінің параметрлерін топталмаған берілгендер бойынша іздеу

У пен Х-тің арасындағы корреляциялық байланыс сызықтық болсын, онда регрессияның сызығының теңдеуі түзу болады.

Бұл түзулердің теңдеулерін іздеу үшін n тәуелсіз сынақ жүргізілген және нәтижесінде n сандар жұбы алынған:

$$(x_1, y_1), \quad (x_2, y_2), \ldots, \quad (x_n, y_n). \tag{9}$$

Оны бас жинақтан алынған кездейсоқ таңдама ретінде қарастырамыз. Онда осы берілгендер бойынша табылған шамалар мен теңдеулер таңдамалық деп аталады.

Қарапайым жағдайда, егер x мәніне y сәйкес келетін болса, онда ізделінді $\bar{y}_x = kx + b$ теңдеуін $Y = kx + b$ түрінде жазуға болады, мұндағы $k -$ регрессия түзу сызығының бұрыштық коэффиценті және ол былай белгіленеді: $k = \rho_{yx}$, ал регрессия теңдеуі мынадай:

$$Y = \rho_{yx} x + b. \tag{10}$$

ρ_{yx} және b параметрлерін, (9) теңдеуі XOY жазықтығында (10) теңдеуіне бар мүмкіндігінше жақындайтындай етіп таңдап аламыз. $Y_i - y_i$, $\left(i = \overline{1, n}\right)$ айырмасын ауытқу деп атаймыз, мұндағы Y_i шамасы (10) теңдеуі бойынша есептелінеді, y_i - (9)-дегі сындық нүкте. ρ_{yx} және b параметрлерін ауытқулар квадраттарының қосындысы минимал болатындай етіп таңдап аламыз (яғни, ең кіші квадраттар әдісі). $F(\rho, b) = \sum_1^n (Y_i - y_i)^2$ функциясын құрамыз немесе $F(\rho, b) = \sum_{i=1}^n \left(\rho_{yx} x_i + b - y_i\right)^2$.

Минимумын табу үшін жүйе құрамыз:

$$\frac{\partial F}{\partial \rho} = 2\sum_1^n (\rho x_i + b - y_i) x_i = 0$$

$$\frac{\partial F}{\partial \rho} = 2\sum_1^n (\rho x_i + b - y_i) = 0$$

және бұл жерден ρ және b -ға тәуелді екі теңдеуден тұратын сызықтық теңдеулер жүйесін аламыз: $\left.\begin{array}{l} \left(\sum x^2\right)\rho + \left(\sum x\right) b = \sum xy \\ \left(\sum x\right)\rho + \sum nb = \sum y \end{array}\right\}$. Бұны шешсек, онда:

$$\rho = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - \left(\sum x\right)^2}; \quad b = \frac{\sum x^2 \sum y - \sum x \sum yx}{n\sum x^2 - \left(\sum x\right)^2}$$

және нәтижесінде ізделінді $y = \rho x + b$ теңдеуін аламыз.

3.11 Корреляцияның таңдамалық коэффициенті

Бірақ та, жаңа шама – корреляция коэффицентін енгізу арқылы регрессия теңдеуін басқа түрде жазу ыңғайлы. $b = \bar{y} - \rho_{yx}\bar{x}$ шамасын регрессия теңдеуіне қойсақ:

$$\bar{y}_x = \rho_{yx}x + \bar{y} - \rho_{yx}\bar{x}$$

$$\bar{y}_x - \bar{y} = \rho_{yx}(x - \bar{x}), \tag{13}$$

(12)-ден $\overline{x^2} - (\bar{x})^2 = \sigma_x^2$ екенін ескере отырып, ρ_{yx}-ті табамыз:

$$\rho_{yx} = \frac{\sum n_{xy}xy - n\bar{x}\bar{y}}{n\left(\overline{x^2} - (\bar{x})^2\right]} = \frac{\sum n_{xy}xy - n\bar{x}\bar{y}}{n\sigma_x^2},$$

$$\rho_{yx} \cdot \frac{\sigma_x}{\sigma_y} = \frac{\sum n_{xy}xy - n\bar{x}\bar{y}}{n\sigma_x^2},$$

r_B – корреляцияның таңдамалық коэффициенті :

$$\rho_{yx} = r_B \cdot \frac{\sigma_y}{\sigma_x}.$$

$\bar{y}_x - \bar{y} = r_B \dfrac{\sigma_y}{\sigma_x}(x - \bar{x})$- Y-тің X-ке байланысты түзу сызықты регрессиялық теңдеуі.

Негізгі әдебиеттер

1 Вентцель Е.С., Овчаров Л.А. Теория вероятностей (задачи и упражнения). М., 1973.
2 Гихман И.И., Скороход А.В., Ядренко М.И. Теория вероятностей и математическая статистика. Киев, 1979.
3 Гмурман В.Е. Руководство к решению задач по теории вероятностей и математической статистике. М., 1998.
4 Гнеденко Б.В. Курс теории вероятности. М., 1969.
5 Емельянов Г.В., Скитович В.П. Задачник по теории вероятностей и математической статистике. Л.,1967.
6 Чудесенко В.Ф. «Сборник заданий по специальным курсам высшей математики», М., 1983

Қосымша әдебиеттер

1. Хисамиев Н.Г., Тыныбекова С.Д., Конырханова А.А. Математика. 2 том, Өскемен, 2006.

2. Коваленко И.Н., Филиппова А.А. Теория вероятностей и математическая статистика. М., 1982.
3. Крамер Г. Математические методы статистики. М., 1975.
4. Мешалкин Л.Д. Сборник задач по теории вероятностей. М., 1963.
5. Пугачев В.С. Теория вероятностей и математическая статистика. М., 1979.
6. Сборник задач по теории вероятностей, математической статистике и теории случайных функций / Под ред. А.А. Свешникова. М., 1965.
7. Севастьянов Б.А., Чистяков В.П., Зубков А.М. Сборник задач по теории вероятностей. М., 1980.
8. Тутубалин В.Н. Теория вероятностей. М., 1972.
9. Чистяков В. П. Курс теории вероятностей. М., 1982.
10. Гмурман В.Е. Теория вероятностей и математическая статистика. М., 1998.