№12 дәріс сабағы

Бас және таңдама жиынтықтар. Таңдама әдістері. Таңдаманың статистикалық үлестірімдері. Полигон және гистограмма.

Анықтама. Бақылау және зерттеу объектілерінің барлық жиынын *бас жинақ* деп айтамыз.

Анықтама. Алынған кез келген нысандар жиынын *таңдама жиынтығы* деп айтамыз.

Анықтама. Таңдама жиынтығындағы (немесе генералды жиынтықтағы) нысандар санын *таңдаманың көлемі* деп айтамыз.

*Анықтама.*Таңдаманың вариациялық қатары деп элементтері шамасы бойынша реттелген, яғни, $x^{(1)}, x^{(2)}, \ldots, x^{n}$ түрінде жазылған, мұндағы $x^{(1)} \le x^{(2)} \le \ldots \le x^{n}$, қатарды айтамыз. Көлемі n болатын таңдамада x_i элементі k_i рет кездессін, онда k_i саны x_i элементінің жиілігі деп аталады және $\sum_i k_i = n.$

Анықтама. Статистикалық қатар деп (x_i, k_i) жұптар тізбегін айтамыз. Көбінде, статистикалық қатар кесте түрінде жазылады, бірінші жолы x-тің мәндерінен тұрады, ал екінші жолы – сәйкес жиіліктен тұрады. Таңдаманың статистикалық үлестірімін, сонымен қатар, аралықтар тізбегі мен оларға сәйкес жиіліктері арқылы беруге болады (осы аралыққа түсетін варианталардың жиіліктерінің қосындысы).

3.3 Үлестірімнің эмпирикалық функциясы

1933 жылы совет математигі В.И. Гливенко математикалық статистиканың негізгі теоремасын дәлелдеді. Бұл теоремада X кездейсоқ шамасының үлестірім функциясын жуықтап алу ережесі көрсетілген. Оның мағынасы мынада: кез келген нақты x саны үшін, $x_1,\ldots,x_n$ таңдамасынан алынған $x_n < x$ теңсіздігін қанағаттандыратын x_n санының қатысты жиілігін $n(x)$ деп белгілейміз. Сонымен, барлық сан түзуінде $n(x)$ функциясы берілген.

Анықтама. $F^*(x) = \dfrac{n(x)}{n}$ функциясы X кездейсоқ шаманың таңдамасының эмпирикалық үлестірім функциясы деп аталады.

Үлестірімнің эмпирикалық функциясының қасиеттері:

1. Эмпирикалық функцияның мәндері: $F^{\bullet}(x) \in [0,1]$.
2. $F^*(x)$ - кемімелі емес функция.
3. Егер $x_{\min}$ - ең кіші варианта болса, ал $x_{\max}$ - ең үлкен варианта болса, онда $F^*(x) = 0$ болады, егер $x < x_{\min}$ және $F^*(x) = 1$ болады, егер $x > x_{\max}$.

3.4 Полигон және гистограмма

Үлестірімнің эмпирикалық функциясынан басқа ықтималдықтың тығыздық функциясының аналогын бейнелеу де пайдалы болады. Бұл екі тәсілмен жүзеге асады. Әрбір x_k үшін n_k жиіліктерін санаймыз. Бұл мәндерді координата жазықтығында бейнелесек, сынық пайда болады. Пайда болған сынық жиіліктер полигоны деп аталады. Бұл график әрбір мәннің қаншалықты жиі кезігетіндігін көрсетеді. Жиіліктің орнына көбінде қатысты жиілік $\frac{n_k}{n}$ алынады және оған сәйкес полигон салынады.

Енді жиілік (қатысты жиілік) гистограммасы деген не, соған тоқталайық. Оны салу үшін, барлық $[x_{\min}, x_{\max}]$ аралығы ұзындықтары h болатын тең бөліктерге бөлінеді. Олардың әрбіреуі үшін осы аралыққа түсетін бақылау мәндерінің саны саналады. Егер Δl_S аралығындағы мәндер саны n_S болса, онда табаны Δl_S және биіктігі $\frac{n_S}{n}$ болатын тік төртбұрыш салынады. Сөйтіп, жиіліктің гистограммасын саламыз. Барлық көпбұрыштардың ауданы барлық бақылау құбылысының санына тең, яғни, таңдаманың көлеміне тең.

Анықтама. Жиіліктің гистограммасы деп табандары ұзындықтары h (бөлік аралықтар ұзындығы), ал биіктіктері - $\frac{h_S}{h}$ (жиілік тығыздығы) болатын тік төртбұрыштардан тұратын баспалдақты фигураны айтамыз. Жиілік гистограммасының ауданы таңдаманың көлеміне тең.

Гистограммадан кездейсоқ шаманың ықтималдығының тығыздығының неге тең екенін байқай аламыз, ал үлестірімнің эмпирикалық функциясынан үлестірімнің теориялық функциясын жуықтап алуға болады.