СТАТИСТИЧЕСКА ОЦЕНКА НА ПОДГОТОВКАТА ПО БЕЗОПАСНОСТ НА ТРУДА НА РАБОТЕЩИТЕ В ПЪТНОТО СТРОИТЕЛСТВО В ПЛАНИНСКИ УСЛОВИЯ

Доц. д-р инж. Анета Георгиева
Технически университет-Варна

Резюме: *В статията са разгледани въпроси, свързани с методите за статистическа оценка на познанията по безопасност на труда на работещи в областта на пътното строителство в планински условия. Извършена е статистическа обработка на данни, получени при независим интелектуален експеримент (анкетно проучване) на лицата от различните групи, представено в [6]. Използван е модел, основаващ се на проверка на хипотезите за еднаквост или различие на разпределенията на частичните съвкупности, получени по резултатите от анкетирането. Предложени и реализирани са два модела, включващ алгоритми за проверка на хипотези за равенство (различие) на две математически очаквания. Чрез тях са решени два основни вида задачи за оценка на познанията по безопасни условия на труд при пътно-строителни дейности в планински условия, а именно проверка на хипотезите за равенство или различие на параметрите и законите на разпределение върху данни, получени след непосредствено анкетиране на участъците от всяка от оценяваните групи (обучение) и проверка на хипотезите за равенство или различие на параметрите и законите на разпределение за случайни величини (извадки) при отсъствие на априорна информация за източника на изходните данни (разпознаване). На базата на представените в статията теоретични разсъждения и практическа реализация на разработените модели са направени констатации и изводи, относно извършената статистическа оценка на познанията и подготовката на работещите в разглеждания сектор и специфичните особености на приложение на предложените модели.*

Ключови думи: *безопасни условия на труд, проверка на хипотези, критерий на Пиърсън, разпределение на Стюдънт, дисперсионно отношение, частични съвкупности*

STATISTICAL ASSESSMENT OF OCCUPATIONAL SAFETY TRAINING OF WORKERS IN ROAD CONSTRUCTION IN MOUNTAIN CONDITIONS

Assoc. prof. eng. Aneta Georgieva
Technical University-Varna

Abstract: *The article discusses issues related to methods for statistical assessment of occupational safety knowledge of workers in the field of road construction in mountainous conditions. Statistical processing of data obtained during an independent intellectual experiment (survey) of individuals from different groups, presented in [6], was carried out. A* *model was used based on testing the hypotheses of equality or difference of the distributions of the partial sets obtained from the survey results. Two models were proposed and implemented, including algorithms for testing hypotheses of equality (difference) of two mathematical expectations. They have solved two main types of tasks for assessing knowledge of safe working conditions in road construction activities in mountainous conditions, namely, testing the hypotheses for equality or difference of the parameters and distribution laws on data obtained after direct surveying of the sections of each of the assessed groups (training) and testing the hypotheses for equality or difference of the parameters and distribution laws for random variables (samples) in the absence of a priori information about the source of the initial data (recognition). Based on the theoretical considerations presented in the article and the practical implementation of the developed models, findings and conclusions have been made regarding the statistical assessment of the knowledge and training of workers in the sector under consideration and the specific features of the application of the proposed models.*

Keywords: *safe working conditions, hypothesis testing, Pearson's test, Student's distribution, dispersion relation, partial sets*

1. Обща постановка

Стратегическата рамка на ЕС за безопасност при работа 2021-2027г. и последвалата я национална програма за сигурност и здраве при работа (2022-2029г.) [6] поставят повишени изисквания за подобряване на превенцията на трудовите злополуки и професионални заболявания в условията на динамично променящата се икономическа среда. Според статистическите данни на Евростат за 2020г. за възникнали трудови злополуки в индустрията и НОИ строителството е един от най-рисковите отрасли по отношение на безопасността на труда [6]. Това с още по-голяма степен се отнася за строителните дейности по изграждането на пътища в планинските райони, където към общите рискови фактори следва да бъдат добавени и специфичните опасности, свързани с геодинамичните гравитационни процеси (свлачища, срутища и др.), които могат да се появят, както преди започване на строителството, така и по време на изпълнение на строителните дейности.

Въпрос от изключително голямо значение в тези сложни условия, определени от големите обеми на строителните дейности при изграждане на пътища и кратките срокове на изпълнение е набирането на добре подготвени и квалифицирани управляващи и изпълнителски кадри. Проблемите произтичащи от тези противоречия придобиват конкретен смисъл в следните направления: [6]

- наемане на общи, неспециализирали в строителството работещи, които не са преминали през задължителните етапи на обучение по безопасни условия на труд (БУТ);
- допускане на подизпълнители с по-ниска за този вид строителство квалификация и с неопитни строителни техници;
- липса на строителен надзор на всички етапи от строителството;
- липса на строително оборудване и механизация, неправилен подбор на оборудването с последици за БУТ;
- лошо планиране на строителната площадка;

- неспазване на графиците за доставка на материали, ниско качество на материалите, неправилно боравене и употреба на инструменти и материалите на строителната площадка;

- проектна документация с неясна спецификация, работни чертежи с ниско качество, промени и забавяне на изготвянето на проектите;

- лошо професионално управление, намеса на държавни, областни и общински лица и други форсмажорни обстоятелства.

От тези констатации се очертават три направления, свързани с нивото на знанията и уменията на личния състав на пътно-строителните фирми, при които работодателите възприемат обучението по безопасни условия на труд (БУТ) като административна тежест, работещите не разбират или не спазват правилата за безопасност, които често са неразбираеми за тях, поради ниската им квалификация, образование и социален статус. Способите за придобиване на знания и умения по БУТ би следвало да се свеждат до следните изисквания и предварителни знания:

- Изисквания за професионална квалификация на работещите в пътното строителство (строителен техник, пътен строител и помощник пътен строител), получена чрез Системата за професионалното образование и обучение, регламентирана от Закон за професионалното образование и обучение (ЗПОО) [6];

- Знания и умения за БУТ на конкретното работно място, които трябва да бъдат получени при инструктажите преди постъпването и в процеса на работа;

- Допълнителни целенасочени знания и умения, получени под формата на „обучение през целия трудов стаж“ в процеса на продължително упражняване на професията (lifelong learning).

1. В Държавните образователни изисквания (ДОИ) за придобиване на квалификацията „пътен строител“ не са включени условията за БУТ при строителство на пътища в планински райони. В списъка на Единни резултати от учене за придобиване на квалификацията „строителен техник“ е включено направление „Здравословни и безопасни условия на труд“, но към изискваните резултати от обучението не се споменават особеностите на строителството и рисковите фактори от поява на геодинамични процеси при строителството на пътища в планински райони. Завършилите специалност „транспортно строителство“ с образователна квалификационна степен бакалавър и магистър притежават по-задълбочени познания относно инженерно-геоложките характеристики на земната основа и възможните геодинамични явления, без да са изучавали подробно опасностите за работещите и превантивните мерки за намаляване на риска при отчитане на специфичните условия на строителство в районите с геодинамична опасност [6].

2. Що се отнася до ефективността на инструктажите при започване на работа и при обсъждания и анализи на сериозни нарушения са направени следните констатации[1]:

- Не са разработвани програми за провеждане на обучението по безопасност;
- Не са определени конкретни видове обучения;
- Не е провеждано обучение по безопасни условия на труд на длъжностните лица, които провеждат инструктажите;
- Липсват писмени данни за проведени инструктажи;

- Не са представени писмени инструкции за безопасно използване на работното оборудване;

- Не са подготвени инструктажи по БУТ при строителство на пътища в високопланински райони;

- Не е предоставена на работещите необходимата информация за рисковете на работното място.

Обобщавайки тези два способа за придобиване на знания и умения по БУТ, стигаме до извода, че при подготовката на изпълнителския и ръководен състав в областта на пътното строителство не се разглеждат специфичните условия на строителство при наличие на геоложки гравитационни процеси, като рискови фактори за безопасността на труда.

3. Знания по безопасни условия на труда, получени чрез съчетаване на сравнително добро базово специализирано образование по геодинамични процеси с умения по охрана на труда, усвоени на базата на собствен опит и организирани форми на обучение (дотолкова, доколкото те съществуват) при продължително упражняване на професията. Такава форма на обучение би трябвало да почива на продължителен трудов стаж и нова, непрекъснато развиваща се система от правила.

В търсенето на пътища за подобряване на системата за осигуряване на безопасни условия на труда си задаваме въпроса за достатъчност на така представените форми на обучение. Отговор на тези въпроси може да бъде намерен след статистическа обработка на данни, получени в процеса на независим интелектуален експеримент, т.е. чрез анкетиране на лицата от различните групи, представено в [6].

В анкетното проучване са включени по 10 представители на ръководния и изпълнителски състав от 14 фирми, ангажирани с високопланинско пътно строителство (140 души). Анкетираните са помолени да изразят мненията си с „да“ или „не“ на 86 въпроса, имащи отношение към познаването на БУТ, разпределени в 18 групи: обща оценка на производствения риск, правила за безопасна работа, организация и провеждане на инструктажите, квалификация на водачите на пътно-строителните машини, осигуряване с предпазни средства и специализирано облекло, организация на безопасна работа с основните видове техника (осем различни вида), електро-обзавеждане, сигнализация и схема на териториална структура на работната площадка, безопасност на работа при различни условия-изкопи, насипи и работа на височина.

Въпросникът за анкетиране по безопасност отговаря на изискванията на ЕС за анкетни проучвания по БУТ [6], като въпросите са зададени така, че всеки положителен отговор (да) да показва познаване или желание за спазване на условията за безопасност на труда.

При обработка на данните е направено разделяне на анкетираните на базата на личната им самооценка на две групи по 70 човека според техния образователен, социален и производствен статус на две групи;

X - работещи в областта на пътното строителство над 10 години със сравнително добра базова професионална подготовка по специалността и с личен опит за условията на безопасна работа. Данните от тази група, наречена „контролна“ би трябвало да определят, съгласно терминологията на изследване на операциите, така наречената еталонна съвкупност (контролни данни).

Y - Участниците в пътно-строителния процес, постъпили на работа след представяне на необходимите документи и преминали през задължителен общ и специализиран инструктаж с трудов стаж под 10 години. Данните от тази група, наречена „експериментална“ са обект на сравнение с еталонните и на разпознаване и прогнозиране на бъдещото им развитие във фирмата след целенасочено продължаващо обучение.

Основна теза: Достатъчно условие за спазването на изискванията за безопасност на труда са: входящите документи, задължителният инструктаж и уменията, придобити по време на трудов стаж с продължителност над 10 години.

В търсенето на ясна и убедителна количествена оценка на така поставената теза ще използваме модел, основаващ се на проверка на хипотезите за еднаквост или различие на разпределенията на частичните съвкупности, получени по резултатите от проведеното анкетиране. Съображенията ни за приемане на този подход произтичат от следното.

В теорията на приемане на решения статистическата хипотеза представлява просто потвърждение на едно или друго свойство на разпределението на изследваната случайна величина, а методът, който се използва за вземане на решения за справедливостта на тази хипотеза се нарича проверка на хипотези. Проверката на статистическа хипотеза се основава на принципа, че малко вероятното събитие се смята практически за невъзможно, а събитието с вероятност близка до единица се приема практически за достоверно [1].

Тези общи разсъждения са достатъчни за използване на методите на статистическа проверка на хипотези като достатъчно достоверни и даващи добра основа за приемане на конкретни управляващи алтернативи, без да изключваме възможностите и на други подходи – например такива като корелационен или дискриминантен анализ [3].

2. Модел за доказване на основната теза за достатъчност на априорните условия за безопасност на труда при пътно строителство в планински условия

Обща схема на модела.

Като следваме основните положения на статистическата теория за проверка на хипотези [2] и смисъла на основната теза, формулираме най-общо нулевата хипотеза така:

H_0 : Априорните условия за приемане на работа (лични документи, входящ инструктаж и личен трудов стаж) са достатъчни условия за спазването на изискванията за безопасност на труда.

Алтернативната хипотеза е:

$H_a = H_1$: Априорните условия за приемане на работа не са достатъчни и е необходимо провеждането на специализирано продължаващо обучение по БУТ с нови практически и дидактически методи.

В една по-конкретна форма с отчитане на две статистически съвкупности (според отговорите на анкетираните с „да“ или „не“), приемането на нулевата хипотеза означава идентичност или близост на законите и параметрите на разпределение на сравняваните извадки, а алтернативната хипотеза се приема, ако законите и параметрите на разпределенията са съществено различаващи се. В една още по-конкретна форма свеждаме задачата за съставянето на модела до следните процедури:

- Приемане на процедури за равенство на параметрите на разпределенията на сравняваните съвкупности.

  - Проверка на хипотезата за равенство на математическите очаквания на двете съвкупности.

  - Проверка на хипотезата за равенство на дисперсиите на двете съвкупности (ако е необходимо)

- Проверка на хипотези за законите на разпределение на статистическите съвкупности.

  - Установяване на закона на разпределение на изследваната случайна величина въз основа на статистически данни при издигане на някаква хипотеза H_o за това разпределение (например нормално), ако това е необходимо.

  - Проверка на хипотеза за еднородност на две разпределения, отговарящи на емпиричните статистически съвкупности от двете групи.

В следващия текст се разглеждат теоретичните модели за приемане на така формулираните хипотези.

2.1. Проверка на хипотези за равенство на параметрите на разпределение на сравняваните съвкупности

2.1.1. Алгоритъм за проверка на хипотези за равенство (различие) на две математически очаквания

1) Област на приложение

Целта на проверката на хипотези е решаване на една от следните задачи:

а) да се провери, различават ли се средните стойности на две случайни величини;

б) да се провери, дали средната стойност на едната от двете случайни величини е по-голяма от средната стойност на другата;

в) да се провери, различават ли се средните стойности на повече от две (например три) случайни величини.

2) Данни

Сравняват се две извадки, взети съответно от генералните съвкупности X, Y с обеми n_1, n_2 .

3) Предположения при формулиране на критериите и хипотези за проверка

а) X, Y са случайни величини с генерални съвкупности, разпределени по нормален закон.

б) Наблюденията от сравняваните извадки са независими.

в) Дисперсиите $\sigma_1{}^2$, $\sigma_2{}^2$ на извадките могат да бъдат известни или да се определят експериментално техните оценки $s_1{}^2$, $s_2{}^2$. В зависимост от това се получават различни правила за проверка на хипотезите.

В конкретното изследване поради малките обеми на извадките ще работим по процедурата за проверка на хипотези при неизвестни дисперсии $\sigma_i{}^2$. Отбелязваме, че ако сравняваните извадки имат голям обем ($n_i > 30$), разпределението на средните аритметични $\overline{x}$, $\overline{y}$ се стреми към нормалното, независимо от разпределението на случайните величини x, y, което се обяснява с централната гранична теорема [1, 2]. Поради това описваният критерий може да бъде използван и за произволно разпределени случайни величини при големи обеми на извадките.

Ако при сравняване на два класа се означи с μ_1 математическото очакване на случайната величина X, а с μ_2 - математическото очакване на Y са възможни следните хипотези:

а) двустранен критерий: H_0: $\mu_1 = \mu_2$; H_1: $\mu_1 \neq \mu_2$;

б) едностранен критерий: H_0: $\mu_1 \leq \mu_2$; H_1: $\mu_1 > \mu_2$;

в) едностранен критерий: H_0: $\mu_1 \geq \mu_2$; H_1: $\mu_1 < \mu_2$.

Едностранен критерий се използва, когато по някакви съображения има основание да се смята, че двете математически очаквания са центрове на две различни разпределения, какъвто е разглежданият тук случай.

2.1.1.1. Процедура (алгоритъм) за проверка на хипотезите за равенство на математическите очаквания при малки обеми на извадките с неизвестни дисперсии $\sigma_1^2 = \sigma_2^{\ 2}$.

Този случай се среща, когато двете извадки са получени при едни и същи условия на работа. Ако съществува колебание относно равенството на дисперсиите, трябва предварително да се провери хипотезата за равенство на дисперсиите.

При предположение за равенство на дисперсиите проверката се извършва по следната процедура:

- Изчисляват се оценките на математическите очаквания

$$\bar{x} = \frac{1}{n_1}\sum_{i=1}^{n_1} x_i \quad ; \quad \bar{y} = \frac{1}{n_2}\sum_{i=1}^{n_2} y_i \tag{1}$$

и дисперсиите:

$$s_1^2 = \frac{1}{n_1-1}\sum_{i=1}^{n_1}(x_i - \bar{x})^2 \quad ; \quad s_2^2 = \frac{1}{n_2-1}\sum_{i=1}^{n_2}(y_i - \bar{y})^2 \tag{2}$$

- Пресмята се нормираното разстояние между разпределенията:

$$\mathrm{T} = \frac{(\bar{x}-\bar{y})}{\sqrt{(n_1-1)S_1^2+(n_2-1)S_2^2}}\sqrt{\frac{n_1.n_2(n_1+n_2-2)}{n_1+n_2}} \tag{3}$$

Доказано е, че ако е вярна нулевата хипотеза, величината T има разпределение на Стюдънт с $\nu = n_1 + n_2 - 2$ степени на свобода.

От таблиците на разпределението на Стюдънт (Табл.4, Приложение [1]) се отчита $t\,(\nu, \frac{\alpha}{2})$ за двустранен критерий или $t\,(\nu, \alpha)$ – за едностранен критерий.

- Проверка на хипотезите се извършва по следните правила:

а) двустранен критерий: нулевата хипотеза ($\mu_1 = \mu_2$) не се отхвърля, ако $|T| < t\left(\nu, \frac{\alpha}{2}\right)$ и се отхвърля (приема се алтернативната ($\mu_1 \neq \mu_2$)), ако $|T| \geq t\left(\nu, \frac{\alpha}{2}\right)$;

б) едностранен критерий при нулева хипотеза H_0: $\mu_1 \leq \mu_2$ (и с алтернативна H_1: $\mu_1 > \mu_2$) не се отхвърля, ако $T < t(\nu, \alpha)$. Ако $T \geq t(\nu, \alpha)$. Нулевата хипотеза се отхвърля и се приема алтернативната H_1: $\mu_1 > \mu_2$;

в) едностранен критерий при нулева хипотеза H_0: $\mu_1 \geq \mu_2$ (с алтернативна H_1: $\mu_1 < \mu_2$) не се отхвърля, ако $T > -t(\nu, \alpha)$. В противен случай, ако $T \leq -t(\nu, \alpha)$ нулевата хипотеза се отхвърля и се приема алтернативната.

2.1.1.2. Дисперсиите $\sigma_1^{\,2}$ и $\sigma_2^{\,2}$ са неизвестни и не са равни ($\sigma_1^{\,2} \neq \sigma_2^{\,2}$).

- Изчисляват се $\bar{x}$, $\bar{y}$ и $s_1^{\,2}$, $s_2^{\,2}$ по формулите (1) и (2).
- Изчислява се величината:

$$\mathrm{T} = \frac{(\bar{x}-\bar{y})}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}} \tag{4}$$

Тази величина има разпределение на Стюдънт с брой на степените на свобода:

$$\nu = \frac{\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1+1}+\frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2+1}} - 2 \tag{5}$$

От таблиците на разпределението на Стюдънт (табл.4, Приложение [1],) се отчита $t\,(\nu, \frac{\alpha}{2})$ за двустранен критерий или $t\,(\nu, \alpha)$ – за едностранен критерий.

- Проверката на хипотезите се извършва по същите правила, както в т.2.1.1.1 по стойности на T и ν, изчислени по формули (4) и (5).

2.1.2. Алгоритъм за проверка на хипотези за равенство на две дисперсии

1) Област на приложение

Тези задачи се формулират, когато е необходимо да се сравняват две методики на изследване или влиянието на различните условия при експериментиране, или в статистическия контрол на качеството или за проверяване на условията за приложимост на други критерии за проверка на хипотези (например при проверка на хипотезата за равенство на две математически очаквания може да се наложи предварително да се провери, дали дисперсиите на двете случайни величини са равни).

2) Данни

Сравняват се две извадки, взети съответно от генералните съвкупности X и Y. Предполага се, че извадката от генералната съвкупност X е с обем n_1 наблюдения, т.е. $\{X\}$: x_1, x_2, ….., x_n, а извадката от генералната съвкупност Y е с n_2 наблюдения, $\{Y\}$: y_1, y_2, ….., y_n,

3) Предположения и хипотези

- X и Y са случайни величини, разпределени по нормален закон.
- Наблюденията от двете извадки са независими.

Ако със σ_1^2 се означи дисперсията на случайната величина X и със σ_2^2 – дисперсията на Y, са възможни следните хипотези:

а) двустранен критерий: H_0: $\sigma_1^{\,2} = \sigma_2^{\,2}$; H_1: $\sigma_1^{\,2} \neq \sigma_2^{\,2}$;

б) едностранен критерий: H_0: $\sigma_1^{\,2} \leq \sigma_2^{\,2}$; H_1: $\sigma_1^{\,2} > \sigma_2^{\,2}$;

в) едностранен критерий: H_0: $\sigma_1^{\,2} \geq \sigma_2^{\,2}$; H_1: $\sigma_1^{\,2} < \sigma_2^{\,2}$.

4) Процедура за проверка. Критерий на Фишер

В началото ще коментираме някои особености на критерия за проверка на хипотезата за равенство на две дисперсии, които дават възможност двете едностранни хипотези да се проверяват лесно чрез едно и също правило.

В основата на процедурата за проверка на хипотези е заложен критерият на Фишер, предложен през 1938 г. [5, 10]

Ако със $s_1^{\ 2}$ се означи оценката на дисперсията $\sigma_1^{\ 2}$, а със $s_2^{\ 2}$ – оценката на дисперсията $\sigma_2^{\ 2}$, като $s_1^2 > s_2^2$ е доказано, че отношението на двете оценки има разпределение на Фишер съответно с $v_1 = n_1 - 1$ и $v_2 = n_2 - 1$ степени на свобода. Това твърдение е в сила, ако с s_1^2 сме означили по-голямата дисперсия $s_1^2 > s_2^2$. Двустранният критерий за проверка на хипотезата (а) е:

$$\frac{1}{F\left(\frac{\alpha}{2},\nu_1,\nu_2\right)} < \frac{s_1^2}{s_2^2} < F\left(\frac{\alpha}{2},\nu_1,\nu_2\right), \tag{6}$$

където $\frac{s_1^2}{s_2^2}$ - дисперсионно отношение.

Нивата на значимост α са избрани в таблиците на разпределение на Фишер равни на 0,05 или 0,01 (Табл.5, Табл.6 от [1]). От тези таблици се вижда, че за тези нива на значимост при произволен брой степени на свобода е изпълнено условието:

$$F\left(\frac{\alpha}{2},\nu_1,\nu_2\right) \geq 1$$

Следователно ако в числителя на дисперсионното отношение винаги се поставя по-голямата от двете оценки (например $s_1^2 > s_2^2$), лявата страна на неравенството (6) винаги се изпълнява, поради което не е нужно да се проверява. Това означава, че трите случая се свеждат до проверка на една хипотеза:

$$H_0: \sigma_1^{\ 2} = \sigma_2^{\ 2} \ \text{ и } \ H_1: \sigma_1^{\ 2} > \sigma_2^{\ 2}; \tag{7}$$

5) Алгоритъм за проверка на хипотези

- Изчисляват се оценките:

- на математическите очаквания

$$\bar{x} = \frac{1}{n_1}\sum_{i=1}^{n_1} x_i \ ; \qquad \bar{y} = \frac{1}{n_2}\sum_{i=1}^{n_2} y_i$$

- на дисперсиите

$$s_1^2 = \frac{1}{n_1-1}\sum_{i=1}^{n_1}(x_i-\bar{x})^2 \ ; \qquad s_2^2 = \frac{1}{n_2-1}\sum_{i=1}^{n_2}(y_i-\bar{y})^2$$

- Съставя се дисперсионното отношение:

$$F = \frac{s_1^2}{s_2^2} \ ,$$

като в числителя се поставя по-голямата от двете оценки (в случая предполагаме, че това е s_1^2).

- Изчислява се броят на степените на свобода за числителя $v_1 = n_1 - 1$ и за знаменателя $v_2 = n_2 - 1$. При зададено ниво на значимост α от таблиците на разпределение на Фишер (табл.5, табл.6 на [1]) се отчита $F(\alpha, \nu_1, \nu_2)$.

- Проверка

Ако $F \leq F(\alpha, \nu_1, \nu_2)$ нулевата хипотеза, H_0: $\sigma_1{}^2 = \sigma_2{}^2$ не се отхвърля.

Ако $F > F(\alpha, \nu_1, \nu_2)$ нулевата хипотеза се отхвърля, а се приема алтернативната H_1: $\sigma_1{}^2 > \sigma_2{}^2$.

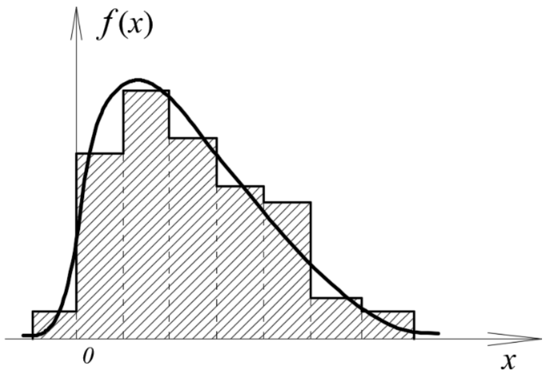
2.2. Проверка на хипотези за законите на разпределение

Представените в предходното изложение методи и подходи за вземане на статистически решения се основаваха на запазването на вида на закона на разпределение на изучаваните случайни явления. В действителност видът на разпределението на дадено случайно явление е хипотетичен модел, който не винаги може да бъде достатъчно убедително обоснован по теоретичен път. Обосноваването на един или друг модел може да стане самo въз основа на резултатите от практиката, т.е. по данните от частичните съвкупности, с които изследователят разполага.

2.2.1. Критерий χ^2 на Пиърсън за установяването на закона на разпределение

Установяването на закона на разпределение на случайната величина X въз основа на статистическите данни се състои в това, че изследователят издига някаква хипотеза, H_0 за вида на това разпределение. Формулирането на хипотезата H_0 обикновено се основава на определени теоретични постановки, на опита и на някаква предварителна експериментална информация. При известни експериментални данни и приета функция (или плътност или хистограма) на разпределение се търси отговор на въпроса за съгласуването на теоретичното и статистическото разпределение [7].

Предполагаме, че статистическото разпределение на фиг.1 е апроксимирано с теоретичната зависимост $f(x)$, като разликите между двете разпределения не могат да бъдат избягнати.



Фиг. 1

Поставя се въпросът, дали тези различия са причинени от случайни обстоятелства, свързани с ограничения обем на наблюденията или те са причинени от това, че теоретичната зависимост не е подходяща за изглаждане на статистическото разпределение. Отговор на този въпрос може да бъде намерен с помощта на статистиката, известна като критерий на съгласие, който служи за оценка на степента на съответствие между издигнатата хипотеза H_0 и статистическите данни [7].

За приемането или отхвърлянето на хипотезата H_0 се разглежда величината U, определяща степента на различие между теоретичното и математическото разпределение. Тя е случайна величина, чийто закон на разпределение зависи от закона на разпределение на случайната величина X и от броя на опитите n. При вярна хипотеза H_0 законът на разпределение на величината U се определя от броя на наблюденията n и от теоретичната функция (Fx) или плътност ($f(x)$) на разпределение. Разглеждаме един предполагаем случай, при който този закон е известен, а избраната мярка на различие U е приела някаква конкретна стойност u. Питаме се, може ли да се обясни това различие на случайни причини или че различието между теоретичното и статистическото разпределение е толкова съществено, че хипотезата H_0 не може да се приеме.

По данните от проведените наблюдения (представляващи частична съвкупност) се изчисляват стойността на критерия на съгласие и вероятността тази стойност да бъде в съответствие с издигнатата хипотеза H_0. Ако тази вероятност $P = 1 - \alpha$ е достатъчно голяма (т.е. вероятността за отхвърляне на H_0 при положение, че тя е вярна е по-малка от някакво зададено ниво на значимост α), хипотезата може да се приеме като не противоречаща на резултатите от наблюдението. В противен случай хипотезата следва да се отхвърли и изследователят може да издигне и провери друга хипотеза или да събере повече статистически данни за повторна проверка на верността на H_0.

Възниква въпросът за начина на избиране на мярката на различие U. Показано е [7], че при някои способи за нейния избор законът на разпределение на величината U има достатъчно прости свойства и при голям брой на наблюденията n, не зависи от функцията $F(x)$. Такава мярка на различието се използва в един от най-широко използваните критерии на съгласие, известен като „критерий χ^2" на Пиърсън.

<u>Критерий на съгласие χ^2 на Пиърсън</u>

1. Област на приложение и предпоставки.

Критерият "χ^2" е един от най-разпространените и универсални методи за проверка на хипотезите за закона на разпределение на случайните величини. Той може да се използва, както за непрекъснати, така и за дискретни случайни величини независимо от закона на тяхното разпределение. При това са възможни две еднакво важни за практиката разновидности:

- Проверка на съответствието на разпределението на емпиричните данни с предложеното теоретично разпределение.
- Проверка на еднородността на разпределението (еднакъв вид на закона на разпределение) на две случайни величини X и Y или на две случайни извадки на една случайна величина.

Критерият χ^2 е свързан с групировка на данните от наблюденията в разреди от вида:

Таблица 1

L_i	$x_1 ; x_2$	$x_2 ; x_3$	……	$x_k ; x_{k+1}$
P_i*	P_1*	P_2*		P_k^n

където P_i* - статистическа честота (вероятност) на наблюденията в разряд №i.

При избран теоретичен закон на разпределението се определят теоретичните вероятности на попадане на случайната величина във всеки от разрядите: $P_1, P_2, \ldots P_к$. При избрана мярка на различие между теоретичните и статистическите разпределения, като сума от квадратите на отклоненията $(P_i^* - P_i)^2$ се получава след сумиране по всички разряди $l = 1 \div k$

$$U = \sum_{i=1}^{k} C_i \, (P_i^* - P_i)^2 \qquad (8)$$

Коефициентите C_i (тегла на разрядите) се въвеждат, тъй като в общия случай отклоненията, отнасящи се към различни разряди не са равностойни по значимост.

Пиърсън е показал, че ако се положи

$$C_i = \frac{n}{P_i} \qquad (9)$$

при голямо n законът на разпределение на величината U има съвсем просто свойство: той практически не зависи от функцията на разпределение $F(x)$ и от броя на опитите n, а зависи само от броя на разредите k, а именно този закон при увеличение на n се приближава към разпределението χ^2, пресмятащо се с гама функции, представени в специални таблици [1, 7].

При този избор на коефициентите C_i мярката на различие, означена като χ^2 се пресмята по зависимостта:

$$\chi^2 = n \sum_{i=1}^{k} \frac{(P_i^* - P_i)^2}{P_i} \qquad (10)$$

За удобство на изчисленията (да не се работи с дробни числа с голям брой на нулите) броят на данните n се въвежда под знака на сумата. Отчитайки при това, че

$$P_i^* = \frac{m_i}{n} \; ,$$

където m_i е броят на наблюденията в i-тия разряд уравнение (10) приема вида:

$$U = \chi^2 = \sum_{i=1}^{k} \frac{(m_i - n \,.\, P_i)^2}{n \,.\, P_i} \qquad (11)$$

Разпределението χ^2 зависи от параметъра r, наречен брой на „степените на свобода“. Броят на „степените на свобода“ r е равен на броя на разрядите минус броя на независимите условия („връзки“), наложени върху честотата P_i^*. Такива условия може да бъдат:

- Сумата от честотите е равна на единица във всички случаи: $\sum_{i=1}^{k} P_i^* = 1$
- Съвпадане на теоретичното и статистическото средни значения (зависи от избраното теоретично разпределение): $\sum_{i=1}^{k} \tilde{x}_i . P_i^* = m_x$
- Съвпадане на теоретичната и статистическата дисперсия (зависи от избраното теоретично разпределение): $\sum_{i=1}^{k} (\tilde{x}_i - m_x^*)^2 . P_i = D_x$

За предварително пресметнатите значения на разпределението χ^2 са съставени специални таблици с входни величини: вероятността P и броя на степените на свобода r. (табл.4 в приложение на [7]).

Разпределението χ^2 дава възможност за оценка на степента на съгласуваност на теоретичното и статистическото разпределение. Ако определената по уравнение (11) стойност на χ^2 е по-малка от табличната $\chi^2_{\alpha,\, v}$ ($\chi^2 < \chi^2_{\alpha,\, v}$), където v е степен на свобода, а $\alpha = 1 - P$ е ниво на значимост, т.е.:

$$P(\chi^2 < \chi^2_{\alpha,\, v}) \geq 1 - \alpha \tag{12}$$

може да се приеме изводът, че данните от наблюденията не противоречат на твърдението на Нулевата хипотеза и тя може да се приеме. При $\chi^2 > \chi^2_{\alpha,\, v}$ хипотезата H_0 следва да се отхвърли. [1]

На тази основа прилагането на критерия χ^2 за оценка на съгласуваността на теоретичното и статистическото разпределение се свежда до следното:

1. Определя се мярката на различие χ^2 по уравнение (11).
2. Определя се броят на степените на свобода r като разлика на броя на разрядите k и броя на наложените връзки s. ($r = k – s$)
3. По r и χ^2 по таблични данни за χ^2 (табл.4, приложение [7]) се определя вероятността $P = 1 – \alpha$, за това, че величината, имаща разпределение χ^2 с r степени на свобода превишава пресметнатото по уравнение (11) значение на χ^2. Ако тази вероятност е много малка, хипотезата се отхвърля като неправдоподобна. Ако вероятността P е относително голяма, е възможно признаването на хипотезата като непротиворечаща на опитните данни. Въпросът, колко малка може да бъде вероятността Р, за да бъде отхвърлена или преразгледана хипотезата, няма точен отговор; той не може да бъде решен по математически съображения [7]. На практика, ако P се получи по-малко от 0,1, се препоръчва повторна проверка на експеримента или използване на друго теоретично разпределение.

3‘. Възможен е и друг вариант, по-близък до процедурите за проверка на хипотези [2, 8], основаващ се на уравнение (12). В този случай по зададените стойности на нивото на значимост $P = (1 – \alpha)$ и степента на свобода v се определя граничната стойност на величината $\chi^2_{\alpha,\, v}$ от табл.4 в [7]. Тя се сравнява с пресметнатата по уравнение (11) статистическа стойност на χ^2.

- При $\chi^2 < \chi^2_{\alpha,\, v}$ може да се приеме, че данните от наблюденията не противоречат на твърдението на Нулевата хипотеза и тя може да бъде приета;
- При $\chi^2 > \chi^2_{\alpha,\, v}$ хипотезата H_0 следва да се отхвърли.

Този вариант се предлага за проверка на хипотези за отклонения на статистическите условия при поява на смущаващи събития с помощта на програмата за електронни таблици Microsoft Office Excel и статистическата функция CHIINV. Поставя се въпросът, дали отклоненията между емпиричните и теоретичните честоти се обясняват със случайни причини (H_0) или тези отклонения, свързани с изследваното събитие са толкова големи, че нулевата хипотеза е непригодна и трябва да се приеме алтернативната H_a. В този софтуерен продукт статистическата оценка на χ^2 запазва вида на уравнение (4), но обекти на сравнение са статистическите (m_i) и теоретичните, пресметнати по софтуера честоти (m_i'):

$$\chi^2 = \sum_{i=1}^{n} \frac{(m_i - m_i')^2}{P_i} \qquad (13)$$

В тази зависимост n е броят на използваните разряди, а m_i и m_i' са съответно брой на възникване на изследваното събитие във всеки разряд.

Така пресметнатото значение на χ^2 се сравнява с табличното $\chi^2_{\alpha,\,\nu}$ и се приема решение съгласно (12).

2.2.2. Проверка на еднородността на разпределенията на две случайни извадки

Поставя се задачата за използване на критерия χ^2 за проверка на хипотезата за еднородност на разпределенията на две случайни величини X и Y или на една и съща случайна величина в две различни условия или времена. За решаването на подобни задачи, възникващи често в практиката, е възможно използването на схемата на приложение на критерия χ^2, свеждаща се до сравнение на статистическата стойност на χ^2 за двете проверявани разпределения с теоретическите таблични значения на случайната величина, разпределена по $\chi^2_{\alpha,\,\nu}$ (табл.4 от [7],).

В резултат на проведените експерименти (наблюдения) са получени две редици данни (частични съвкупности) $x_1, x_2 \dots x_{n'}$ и $y_1, y_2 \dots y_{n'}$ с обеми n' и n'', разпределени в еднакъв брой (l) разреди. Не е задължително разрядите да са едни и същи или да са с еднакви стъпки.

Абсолютните честоти за всяка случайна величина n_i' и n_i'' са свързани със съотношенията:

$$\sum_{i=1}^{l} n_i' = n' \;;\quad \sum_{i=1}^{l} n_i'' = n'' \qquad (14)$$

Проверката на нулевата хипотеза H_0, че разпределенията на двете генерални съвкупности, от които са взети частичните съвкупности са еднакви, се извършва с помощта на статистиката (критерия χ^2) [2]:

$$\chi^2 = n'.n'' \sum_{i=1}^{l} \frac{1}{n_i' + n_i''} \left[\frac{n_i'}{n'} - \frac{n_i''}{n''}\right]^2 \qquad (15)$$

В [8, 9] е показано, че величината χ^2, пресмятана по това уравнение при голям брой на данните n_i' и n_i'' има асимптотично χ^2 – разпределение с брой на степените на свобода $\nu = l - 2$, тъй като величините в него са свързани с две линейни зависимости от вида (14).

Проверката на Нулевата хипотеза за еднаквост на двете генерални разпределения, от които са взети частичните съвкупности X и Y се извършва по схемата:

1. Пресмята се мярката на различие между двете различни съвкупности X и Y с помощта на статистиката (критерия χ^2) (15).

2. От табл.4 в [7] се определя граничната стойност на величината $\chi^2_{\alpha,\,v}$ за степен на свобода $v = l - 2$ и вероятност $P = 1 - \alpha$ (при $\alpha = 0.05$).

3. Сравнява се статистиката χ^2, пресметната по уравнение (15) с граничната стойност $\chi^2_{\alpha,\,v}$.

- При $\chi^2 < \chi^2_{\alpha,v}$ – може да се приеме, че данните от наблюденията не противоречат на Нулевата хипотеза за еднаквост на разпределенията и тя може да бъде приета.
- При $\chi^2 > \chi^2_{\alpha,\,v}$ – Нулевата хипотеза следва да се отхвърли.

3. Резултати

С помощта на така предложения модел са решени следните задачи за оценка и приемане на безопасните условия на труд при пътно-строителните дейности в планински условия.

3.1. Проверка на хипотезите за равенство или различие на параметрите и законите на разпределение върху данни, получени след непосредствено анкетиране на участниците от всяка от оценяваните групи.

Този начин на обработка на статистически данни, получени от участниците от двете оценявани групи („белязани" данни) има за цел да определи началната различимост между сравняваните статистически съвкупности и е близък по смисъл до термина „обучение", използван в теорията на приемане на решения [4].

3.1.1. Процедура за проверка на хипотезите за равенство на математическите очаквания с неизвестни дисперсии

Тъй като процедурите за проверка на хипотези за равенство на математическите очаквания (т.2.1.1) са различни в зависимост от равенството на дисперсиите (т.2.1.1.1 при $\sigma_x^{\,2} \neq \sigma_y^{\,2}$) е необходимо да бъде проверена хипотезата за равенство на двете дисперсии, за да бъде избран актуалният алгоритъм.

За целта като използваме изходните данни (колони 1 и 3 на фиг.2) проверяваме хипотезата за равенство на две дисперсии по процедурата на Фишер (т.2.1.2).

<u>Нулева хипотеза H_o :</u> $\sigma_x^{\,2} = \sigma_y^{\,2}$

$$\bar{x} = \frac{1}{n_1}\sum_{i=1}^{n_1} x_i = 52{,}06\ ; \quad \bar{y} = \frac{1}{n_2}\sum_{i=1}^{n_2} y_i = 39{,}61$$

$$s_x^2 = \frac{1}{n_1-1}\sum_{i=1}^{n_1} (x_i - \bar{x})^2 = 228{,}99\ ; \quad s_y^2 = \frac{1}{n_2-1}\sum_{i=1}^{n_2} (y_i - \bar{y})^2 = 253{,}43$$

Означаваме $s_x^{\,2} = s_2^{\,2}$; $s_y^{\,2} = s_1^{\,2}$

Дисперсионно отношение

$$F = \frac{s_1^2}{s_2^2} = \frac{253{,}43}{228{,}99} = 1{,}107$$

От табл.6 [2]

$F = (\alpha; \nu_1; \nu_2) = F\,(0{,}01;\ 17;\ 17) = 3{,}20$

$F = 1{,}107 < F\,(0{,}01;\ 17;\ 17) = 3{,}20$

Нулевата хипотеза за равенство на дисперсиите не се отхвърля.

Приема се: $\sigma_x{}^2 = \sigma_y{}^2$.

Проверката на хипотези за равенство на математическите очаквания се извършва по алгоритъма 2.1.1, чиято пълна реализация е представена на фиг.2.

X	12	57	68	68	67	38	39	61	57	51	57	48	68	44	60	59	28	55
Y	9	42	59	66	61	22	22	43	44	32	44	30	59	21	38	52	28	41

$\overline{X}$	52.06
∑(xi - $\overline{x}$)^2	3892.94
n_1	18
n_1-1	17
s1^2	228.99
s_1	15.13

$\overline{Y}$	39.61
∑(yi - $\overline{y}$)^2	4308.28
n_2	18
n_2-1	17
s_2^2	253.43
s_2	15.92

T	2.4
$\overline{X}$ - $\overline{Y}$	12.44
α = 1 - P	0.1

a) Двустранен критерий		
H_0:	$\mu_x = \mu_y$	|T|≤t
H_a:	$\mu_x \neq \mu_y$	|T|>t
α/2	$\nu = n_1+n_2 - 2$	P
0.05	34	0.9
t		**|T|**
1.69	**<**	**2.4**
H_0 се отхвърля		
Приема се H_a: $\mu_x \neq \mu_y$		

b) Едностранен критерий		
H0:	$\mu_x \leq \mu_y$	T≤t
Ha:	$\mu_x > \mu_y$	T>t
α = 1 - P	$\nu = n_1+n_2 - 2$	P
0.1	34	0.9
t		**T**
1.306	**<**	**2.4**
H_0 се отхвърля:		
Приема се H_a: $\mu_x > \mu_y$		

Фиг.2. Алгоритъм и резултати за проверка на хипотези за равенство на математическите очаквания – обучение

Окончателни резултати.

От $T = 2{,}4 > t\ (34;\ 0.05) = 1.69$ нулевата хипотеза *Ho:* $\mu_x = \mu_y$ се отхвърля. Приема се алтернативната хипотеза *Ha:* $\mu_x \neq \mu_y$

От $T = 2{,}4 > t\ (34;\ 0.1) = 1.306$ нулевата хипотеза *Ho*: $\mu_x \leq \mu_y$ се отхвърля. Приема се *Ha:* $\mu_x > \mu_y$

Валидност на алтернативната хипотеза (табл.1 от [2]) *Ha:* $\mu_x > \mu_y$ се изпълнява (не се отхвърля) при $0{,}55 < P < 0{,}99$, $0{,}01 < \alpha < 0{,}45$

3.1.2. Процедура за проверка на хипотези за еднаквост на разпределенията на статистическите съвкупности, получени след непосредствено анкетиране.

Задачата за проверка на еднородността на разпределението на две случайни величини *X* и *Y* или на две случайни извадки от една случайна величина решаваме по процедурата в т.2.2, основаваща се на критерия на съгласие на Пиърсън χ^2 [2.7]. За целта преобразуваме резултатите от анкетирането (колони 1 и 4 на фиг.2) в два вариационни реда $x_1 \leq x_2 \leq \ldots\ldots \leq x_n$; $y_1 \leq y_2 \leq \ldots\ldots \leq y_n$ от по седем разряда (табл.2).

Таблица 2

Брой (честоти) на попадане на случайните величини

Разреди / Брой положителни отговори	0-10	11-20	21-30	31-40	41-50	51-60	61-70	$\Sigma n'/\Sigma n''$
Целева група над 10г. x_i (n′)	1	0	5	2	5	3	2	18
Целева група под 10г. y_i (n″)	0	1	1	2	2	7	5	18

Нулева хипотеза H_o: (Разпределенията на двете генерални съвкупности, от които са взети частичните съвкупности $\{X\}$ и $\{Y\}$ са еднакви – $Fx = Fy$).

Проверка на хипотезите

Критерий на Пиърсън

$$\chi^2 = n'.n'' \sum_{i=1}^{7} \left[\frac{n_i'}{n'} - \frac{n_i''}{n''}\right]^2 = 8{,}84$$

За $v = l - 2 = 5$ и вероятност $P =$ 0,95 (при $\alpha = 0{,}05$) отчитаме от табл.4 в [7] граничната стойностна величината $\chi^2_{\alpha,\, v}$ =1.145.

Проверка: $\chi^2 = 8.84 > \chi^2_{\alpha,\, v} = 1.145$. Нулевата хипотеза се отхвърля. Приема се алтернативната $Fx \neq Fy$.

Извод: Алтернативната хипотеза за различие на двете разпределения $Fx \neq Fy$ е в сила (не се отхвърля) в интервала $0{,}554 < \chi^2 < 9{,}24$ с вероятности $0{,}10 < P < 0{,}99$, където попада пресметнатата стойност на критерия на Пиърсън.

3.2. Проверка на хипотезите за равенство или различие на параметрите и законите на разпределение за случайни величини (извадки) при отсъствие на априорна информация за източника на информацията – разпознаване.

Поставяме си задачата за разпознаване и прогнозиране на познанията по безопасни условия на труда по информация, получена в резултат на общи наблюдения или инцидентни анализи, като една обща статистическа съвкупност, т.е. при невъзможност за априорно формиране на характеристиките на предполагаемите източници на групова информация. Такава задача е известна в теорията на разпознаване на образи и приемане на оптимални статистически решения, като „самообучение или обучение с учител“, но за нейното решаване в многомерен или едномерен подход е необходимо по-специално отношение към характеристиките на входящата информация.

Като имаме предвид вече разработения в т.II теоретичен апарат и получените в т.III-1 резултати, използваме за апостериорно разпознаване добре отработения апарат за проверка на статистически хипотези, въпреки че при него решенията се приемат по вероятностен критерий, а методите на разпознаване на образи позволяват определянето на броя на верните или грешни класифициращи решения [4].

3.2.1. Процедури за проверка на хипотези за равенство на математическите очаквания – разпознаване на групите по статистическите характеристики по косвени признаци.

При формирането на данните за проверка на хипотезите за равенствата (различията) на разпределенията изхождаме от вече установения факт в т.3.1.1 за широкия вероятностен диапазон на положителните отговори ($0{,}55 < P < 0{,}99$), който се дължи предимно на анкетираните от първата целева група. Поради това, като косвен признак за разпознаване на всяка от целевите групи използваме броя на положителните отговори над определена граница (напр. 60%). Така получените статистически съвкупности за проверка на хипотези са представени на фиг.3 (колони 1 и 3).

Ред на пресмятане:

А. Процедура за проверка на хипотезата за равенство на дисперсиите на двете статистически съвкупности.

Нулева хипотеза H_o: $\sigma_x{}^2 = \sigma_y{}^2$

$$\bar{x} = \frac{1}{n_1}\sum_{i=1}^{n_1} x_i = 2{,}61 \; ; \quad \bar{y} = \frac{1}{n_2}\sum_{i=1}^{n_2} y_i = 2{,}17$$

$$s_x^2 = 2{,}487 \; ; \quad s_y^2 = 1{,}676$$

Означаваме $s_x{}^2 = s_1{}^2$; $s_y{}^2 = s_2{}^2$

Дисперсионно отношение

$$F = \frac{s_1^2}{s_2^2} = \frac{2{,}487}{1{,}676} = 1{,}484$$

От табл.6 [2] $F = (\alpha;\, \nu_1;\, \nu_2) = F\,(0{,}01;\, 17;\, 17) = 3{,}20$

Б. Решение.

$F = 1{,}484 < F\,(0{,}01;\, 17;\, 17) = 3{,}20$

Нулевата хипотеза за равенство на дисперсиите не се отхвърля.

Приема се: $\sigma_x{}^2 = \sigma_y{}^2$.

Проверките на хипотези за равенство на математическите очаквания се извършва по алгоритъма т.2.1.1, чиято пълна реализация е представена на фиг.3.

Окончателни резултати

От $T = 0{,}92 > t\,(34;\, 0{,}1) = 0{,}853$ => Нулевата хипотеза H_o: $\mu_x = \mu_y$ се отхвърля. Приема се H_a: $\mu_x \neq \mu_y$

От $T = 0{,}92 > t\,(34;\, 0{,}2) = 0{,}853$ => Нулевата хипотеза H_o: $\mu_x < \mu_y$ се отхвърля. Приема се H_a: $\mu_x > \mu_y$.

Валидността на алтернативната хипотеза H_a: $\mu_x > \mu_y$ се изпълнява, не се отхвърля при $0{,}55 < P < 0{,}99$, $0{,}01 < \alpha < 0{,}85$.

X	0	0	1	2	2	6	3	2	2	4	4	2	2	4	5	3	3	2
Y	5	4	2	0	3	4	3	2	2	1	1	1	1	2	2	2	1	3

$\overline{X}$	2.61
$\sum(x_i - \overline{x})^2$	42.28
n_1	18
n_1-1	17
s_1^2	2.487
s_1	1.577

$\overline{Y}$	2.17
$\sum(y_i - \overline{y})^2$	28.5
n_2	18
n_2-1	17
s_2^2	1.676
s_2	1.295

T	0.92
$\overline{X} - \overline{Y}$	0.44
α = 1 - P	0.2

a) Двустранен критерий		
H_0:	μx = μy	|T|≤t
H_a:	μx ≠ μy	|T|>t
α/2	$\nu = n_1+n_2 - 2$	P
0.1	34	0.9
t		**|T|**
0.9	**<**	**0.92**
H_0 се отхвърля		
Приема се H_a: $\mu_x \neq \mu_y$		

b) Едностранен критерий		
H_0:	$\mu_x \leq \mu_y$	T≤t
H_a:	$\mu_x > \mu_y$	T>t
α = 1 - P	$\nu = n_1+n_2 - 2$	P
0.1	34	0.9
t		**T**
0.853	**<**	**0.92**
H_0 се отхвърля:		
Приема се H_a: $\mu_x > \mu_y$		

Фиг.3. Алгоритъм и резултати за проверка на хипотези за равенство на математическите очаквания - разпознаване

3.2.2. Процедура за проверка на хипотези за еднаквост на разпределенията на статистически съвкупности, формирани от положителните отговори на задаваните въпроси – разпознаване

Преобразуваме изходните данни (колони 1 и 4 на фиг.3) в два вариационни реда във вид, удобен за прилагане на критерия на съгласие на Пиърсън (табл.3):

$x_1 < x_2 < \ldots\ldots < x_n$; $y_1 < y_2 < \ldots\ldots < y_n$

Таблица 3

Разреди / Брой положителни отговори	0-0,999	1-1,999	2-2,999	3-3,999	4-4,999	5-5,999	6-6,99	Σn′/ Σn″
Целева група над 10г. x_i (n′)	2	1	7	3	3	1	1	18
Целева група под 10г. y_i (n″)	1	5	6	3	2	1	0	18

Нулева хипотеза *H_o*: (Разпределенията на двете генерални съвкупности, от които са взети частичните съвкупности $\{X\}$ и $\{Y\}$ са еднакви – $F_x = F_y$).

Проверка на хипотезата

Критерий на Пиърсън $\chi^2 = 4{,}28$

За $\nu = l - 2 = 5$ и $P = 0,95$ (при $\alpha = 0,05$) определяме граничната стойност на $\chi^2_{\alpha,\nu} = 1.145$. (табл.4 [1]).

Проверка: $\chi^2 = 4.28 > \chi^2_{\alpha,\nu} = 1.145$. Нулевата хипотеза се отхвърля. Приема се алтернативната $Fx \neq Fy$.

<u>Извод</u>: Алтернативната хипотеза за различие на двете разпределения $F_x \neq F_y$ е в сила (не се отхвърля) в интервала $0,554 < \chi^2 < 4,35$ с вероятности $0,50 < P < 0,99$, в който се намира пресметнатата стойност на критерия на Пиърсън.

Заключение

На базата на направените теоретични разсъждения и практическа реализация на съставените модели са направени следните констатации и изводи:

1. Резултатите, получени при реализацията на модела на обучение (т.3.1) позволиха да бъдат доказани хипотезите за различие на характеристиките и законите на разпределение на данните от изследваните целеви групи и на по-добрата подготовка и готовност за изпълнение на изискванията за безопасен труд от групата, получила знания и умения в един продължителен период на непрекъснато обучение.
2. Резултатите, получени при реализацията на модела за разпознаване на групите по косвени признаци (т.3.2) позволяват да бъдат класифицирани двете целеви групи по тяхното отношение към изискванията за БУТ. По този начин за групата, идентифицирана като по-слабо подготвена следва да бъдат предложени нови способи за получаване на знания, близки до идеята за „обучение през целия производствен процес“.

Литература

1. Божанов Е., И.Вучков. Статистически решения в производството и научните изследвания. „Техника“, С.1979 г., 410с.
2. Кендал М.Дж., А.Стюарт. Статистические выводы и связи. Наука. М. 1973.
3. Недев А., М. Консулова-Бакалова, Хр. Ненов, “Нов подход за визуализация, оценка и прогнозиране на състоянието на сложни обекти (с приложение в енергетиката и опазване на околната среда)”, Енергиен форум 17-20 юни 2009, сборник доклади, 455÷461стр.
4. Недев А., Разпознаване на образи и оптимално стохастическо управление. Книга I-ва И.К.Геа-Принт. Варна. 2012 г., 345с.
5. Фишер Р.А. Статистические методи для исследователей. Госстатиздат М.1958
6. Цекова М., Дисертационен труд „Изследване условията за безопасност при строителството на пътища във високопланински райони“, 2023.
7. Вентцель Е.С. Теория вероятностей. „Наука“. М.1968, 572с.
8. Крамер Г., Матрматические методы статистики. „Мир“. М.1965
9. Уилкс С. Математическая статистика. „Наука“. М.1976
10. Fisher.R.A., E.Yates, Statistical tables from biological agricultural and medical research. Oliver and Boyd. Edinburgh. 1938.
11. Naskova P., A.Georgieva. Statistical Assessments of Agrochemical Soil Analysis Results. Journal of Mountain Agriculture on the Balkans (JMAB), Troyan, 2024, Vol. 27, (4), pp. 605–628, ISSN 1311-0489 (Print); ISSN 2367-8364 (Online)