

СТАТИСТИЧЕСКИ ПОДХОД ЗА ОЦЕНКА НА ЕКОЛОГИЧНО ОПАСНИТЕ ПОКАЗАТЕЛИ НА СЪСТОЯНИЕТО ОТ МОРСКАТА ТРАНСПОРТНА ДЕЙНОСТ

проф. д-тн. инж. Асен Недев, ВВМУ „Н.Й. Вапцаров“, Варна
доц. д-р инж. Анета Георгиева, ТУ-Варна,
Александър Близнаков, ВВМУ „Н.Й. Вапцаров“, Варна, докторант

Резюме: *В студията се разглеждат въпроси, свързани с оценка на екологично опасните показатели при извършване на морската транспортна дейност при конкретен технологичен процес, свързан с операции при разтоварването на фосфорит от кораб за насипни товари на пристанищен терминал. Направено е описание на технологичния процес, работните места, функциите на лицата по контрол и изпълнение. На базата на предварително получени експериментални данни, изложен в [3] е разработен математически модел на технологичните и контролни дейности при наличие на два или повече класа на постепенно влошаваща се околна среда. Определени са статистически обосновани граници между последователно влошаващите се класове на състоянието с помощта на методите за статистическо оценяване и проверка на статистически хипотези. За всяка от извършените видове проверки на хипотези е направен анализ на получените резултати. В заключение е посочено, че за получаването на надеждни резултати, които да бъдат основа за бъдещи стандарти за управление на качеството, е необходимо провеждането на специализирани изпитания в широк мащаб от външни и управляващи условия. На базата на такава статистическа основа и с използването на посочените тук теоретични подходи може да се пристъпи към коригиране на стандартите за контрол на качеството.*

Ключови думи: *екологично опасни показатели, товаро-разтоварни корабни дейности, статистически оценки, проверка на хипотези.*

STATISTICAL APPROACH FOR ASSESSMENT OF ENVIRONMENTALLY HAZARDOUS INDICATORS OF THE STATE OF MARITIME TRANSPORT ACTIVITIES

Prof. Dr. Eng. Asen Nedev, Higher Naval Academy, Varna
Assoc. Prof. Dr. Eng. Aneta Georgieva, Technical University of Varna,
Alexander Bliznakov, Higher Naval Academy, Varna, PhD student

Abstract: *The study examines issues related to the assessment of environmentally hazardous indicators when carrying out maritime transport activities in a specific technological process related to operations during the unloading of phosphorite from a bulk cargo ship at a port terminal. A description of the technological process, workplaces, functions of control and*

execution personnel is made. Based on previously obtained experimental data, presented in [3], a mathematical model of technological and control activities in the presence of two or more classes of gradually deteriorating environment has been developed. Statistically justified boundaries between successively deteriorating classes of the condition have been determined using methods for statistical assessment and verification of statistical hypotheses. An analysis of the results obtained has been made for each of the types of hypothesis testing performed. In conclusion, it is indicated that in order to obtain reliable results that will be the basis for future quality management standards, it is necessary to conduct specialized tests on a wide scale of external and control conditions. Based on such a statistical basis and using the theoretical approaches mentioned here, one can proceed to adjust the quality control standards.

Keywords: *environmentally hazardous indicators, loading and unloading ship activities, statistical assessments, hypothesis testing.*

I. Обща постановка

Задачите за висока екологична ефективност на морските транспортни дейности търсят своите адекватни решения в ред общи изисквания и новопоявяващи се регламенти на браншовите и контролиращите организации. Пределните стойности на емисиите на конкретните показатели (на брой над сто) все още не са нормативно решени, независимо от общите постановки на Международната морска организация (ИМО) и рамковите конвенции на ООН за изменение на климата (UNFCCC). Така например в приложенията към конвенцията MARPOL се препоръчва преминаване от пожелателен към задължителен режим на контрол, като се има предвид, че степента на опасност от различните замърсители и вредите, които те могат да нанесат, би следвало да бъдат решавани чрез опити за създаване на регламенти и стандарти за качеството на средата (морска и атмосферна). Основните трудности при въвеждане на пределните нива на емисиите произтичат от общото изискване те да се определят преди всичко от позицията на човешкото здраве, което приведено към концентрациите на вредните вещества означава, че решенията трябва да се вземат в условия на силна размитост и неопределеност. Това ни насочва към идеята, че пределните стойности на конкретните показатели биха могли да бъдат регламентираны не чрез административни пожелания и решения, а след провеждане на многостранна и продължителна статистическа обработка. При това трябва да се има предвид и фактът, че пределните стойности на емисиите могат да бъдат различни в зависимост от вида на рецептора-човек, животно, растение и т.н.

В настоящото изследване насочваме вниманието си към един пренебрегван и слабо изучен проблем, свързан с определянето на емисиите в атмосферния въздух на вредни вещества, породени от голямото разнообразие на технологични процеси, извършвани в обслуждането на морските транспортни средства в брегови условия.

II. Математичен модел за определяне на статистически обосновани граници между последователно влошаващи се класове на състоянието при едномерни и многомерни вектори на наблюдение – обща структура

Предварително отбелязваме, че нашият стремеж е да разработим достатъчно широк абстрактен математически модел, който да позволява определянето на границите между класовете на състояние (добър, междинен, недобър, опасен) на всеки от елементите и съединенията от постоянния и променлив състав и на замърсителите на атмосферния въздух [1, 2]. По-конкретно става дума за определяне на концентрацията на фини прахови частици във въздуха, възникващи по време на товаро-разтоварни процеси на корабите с товари, които са потенциално опасни за здравето на обслужващия персонал. При това изхождаме от предварително полученият експериментален резултат, изложен в [3].

Обща схема на модела

2.1. Числени характеристики и статистически оценки на измервания показател на състоянието.

2.2. Статистическа проверка на хипотези за параметрите на разпределение на измервания екологичен показател.

2.3. Определяне на оптималните гранични стойности на контролирания екологичен показател между класовете на състояние.

III. Числени характеристики и статистически оценки на измервания екологичен показател

За осъществяване на основната (начална) част на изграждането на модела са необходими числените характеристики на положението и разсейването на случайната величина на резултатите от измерванията, т.е. на генералната съвкупност (разположени в лявата страна на табл.1.) и на техните статистически оценки (разположени в дясната страна на табл.1.). Предварително отбелязваме, че изчерпателният анализ на цялата генерална съвкупност обикновено е много сложно, скъпо струващо, а в повечето случаи и практически неосъществимо мероприятие, още повече, че в много случаи (например при увреждания или унищожаване на обекта) то е лишено изобщо от смисъл. Поради това използването на частичните съвкупности (а от тук и на методите на математическата статистика) има не само теоретично значение, но и огромна практическа стойност [4]. Както се вижда от таблица 1 статистическите оценки на положението и разсейването имат математически вид, аналогичен на формулите, по които се определят съответните числени характеристики на дискретните случайни величини.

Табл.1. Математическа основа при определяне на генерални и частични съвкупности

Числови характеристики на случайните величини (генерални съвкупности)	Статистически оценки на числовите характеристики (частични съвкупности)
<p><u>1.Характеристики на положението</u> Математическо очакване (средна стойност, център на разпределението):</p> $M\{X\} = \mu_x = \begin{cases} \text{дискретни величини} \\ \sum_i x_i P(x_i) \\ \text{непрекъснати величини} \\ \int_{-\infty}^{+\infty} x f(x) dx \end{cases}$	<p><u>1.Оценки на положението – статистики</u> Начален емпиричен момент от първи род (средноаритметично на извадката):</p> $\bar{x} = \alpha_1 = \frac{1}{n} \sum_{i=1}^n x_i \approx \frac{1}{n} \sum_{i=1}^k n_i \cdot x_{i,0} \quad (1)$
<p><u>2.Характеристики на разсейването</u> Дисперсия</p> $D\{X\} = \sigma_x^2 = \begin{cases} \text{дискретни} \\ \sum_i (x_i - \mu_x)^2 \cdot P(x_i) \\ \text{непрекъснати} \\ \int_{-\infty}^{+\infty} (x_i - \mu_x)^2 f(x) dx \end{cases}$ $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} \quad (3)$	<p><u>2.Оценки на разсейването</u> Втори емпиричен централен момент – оценка на дисперсията:</p> $m_2 = S_1^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (2')$ <p>или $S^2 = \frac{n}{n-1} \cdot S_1^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (2'')$</p> <p>неизместена ефективна оценка</p>
<p><u>3.Числови характеристики на система от две случайни величини (случаен вектор – X,Y)</u> Корелационен момент (коэффициент на ковариация):</p> $K_{i,j} = M\{(x_i - \mu_{x_i})(x_j - \mu_{x_j})\}, \quad i \neq j$ <p>Коефициент на корелация: $\rho_{i,j} = \frac{K_{i,j}}{\sigma_i \sigma_j}$</p>	<p><u>3. Оценка на числовите характеристики на система от две случайни величини</u> Втори смесен централен момент (корелационен момент по статистическа съвкупност):</p> $K_{xy} = \sum_i \sum_j (x_i - \mu_{x_i})(x_j - \mu_{x_j}) P_{i,j} \quad (4)$ <p>Оценка на коефициента на корелация:</p> $r_{xy} = \frac{K_{xy}}{s_x s_y}$
<p><u>4.Числови характеристики на система от n случайни величини (x₁, x₂, x_n)</u></p> <p>4.1. n математически очаквания $\mu_1, \mu_2, \dots, \mu_n$ 4.2. n дисперсии $D_1 = K_{11}, D_2 = K_{22}, \dots, D_n = K_{nn}$ 4.3. n(n-1) корелационни моменти: $K_{i,j} = M\{(x_i - \mu_i)(x_j - \mu_j)\}, \quad i \neq j \quad (5)$</p> <p>или ковариационна матрица:</p> $V = \begin{bmatrix} K_{11} & K_{12} & \dots & K_{1n} \\ K_{21} & K_{22} & \dots & K_{2n} \\ \dots & \dots & \dots & \dots \\ K_{n1} & K_{n2} & \dots & K_{nn} \end{bmatrix} \quad (6)$ <p>Нормално многомерно разпределение:</p> $f(x) = (2n)^{-\frac{n}{2}} \cdot V ^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\bar{x} - \bar{\mu}_x)^T \cdot V^{-\frac{1}{2}}(\bar{x} - \bar{\mu}_x)\right\} \quad (7)$ <p>където $\bar{x} = (x_1, x_2, \dots, x_n)$ – случаен вектор $\bar{\mu}_x$ – вектор на математическите очаквания</p>	

Основното предназначение на статистическите оценки, предоставени в дясната страна на Табл.1 е да дават възможност за най-точно и достоверно определяне на числените характеристики на генералната съвкупност по данните на ограничени извадки. Поради това тези оценки трябва да са с определени качества, които да позволяват да се използва напълноценно информацията за генералната съвкупност, която се съдържа в дадена извадка или серия от еднакви по обем извадки. Основните свойства, които характеризират статистическите оценки са състоятелност, неизместеност, ефективност и достатъчност [5, 6].

Състоятелност на дадена оценка θ е свойството да се увеличава точността ѝ при увеличение на броя n на наблюденията. Това означава, че математическото очакване на оценката се стреми неограничено към истинската стойност θ при увеличение на n , като същевременно дисперсията ѝ намалява без ограничение. В [7] е доказано, че оценката \bar{x} на математическото очакване, пресметната по уравнение (1) от Табл.1 е състоятелна, а статистиката S_1^2 по (2') представлява състоятелна оценка на дисперсията на генералната съвкупност.

Неизместеност е свойство, близко по смисъл до състоятелността, но то се отнася до поведението на оценките при краен брой на наблюденията n . Оценката се нарича неизместена (без позиционна грешка), ако при всички възможни обеми n на частичната съвкупност математическото очакване на статистиката $\hat{\theta}_n$ е равно на оценяваната характеристика θ , т.е., ако $M\{\hat{\theta}_n\} = \theta$.

Установено е [5, 7], че за статистиката (S_1^2) , пресметната по уравнение (2') от Табл.1 е в сила съотношението:

$$M\{S_1^2\} = \frac{n-1}{n} \cdot \sigma^2 = \sigma^2 - \frac{\sigma^2}{n} \quad (8)$$

S_1^2 е отрицателно изместена статистика на дисперсията σ^2 (изместването е σ^2/n). Поради това в математическата статистика вместо нея се използва статистиката S^2 , пресметната по уравнение (2'') от Табл.1:

$$S^2 = \frac{n}{n-1} \cdot S_1^2 = \frac{n}{n-1} \cdot \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (9)$$

Тази статистика не е нито дисперсия на частичната съвкупност, нито на генералната съвкупност, но е едновременно състоятелна и неизместена оценка на дисперсията на генералната съвкупност при произволен закон на разпределения, което я прави препоръчително използвана [5, 7].

Ефективността на дадена оценка $\hat{\theta}_n$ характеризира вида и свойствата на нейното разпределение, по-точно степента му на групираност около оценявания параметър θ . Сред статистиките, които могат да се използват за оценка на математическото очакване най-малка дисперсия има средноаритметичната стойност \bar{x} , пресметната по уравнение (1) (Табл.1) и в този смисъл тя е ефективна оценка на математическото очакване [8]. Поради това тази оценка е получила изключителен приоритет пред другите възможни оценки.

Ефективна оценка на дисперсията представлява и статистиката S^2 , пресметната по уравнение (9) [5, 7].

Относителната ефективност на две оценки $\hat{\theta}_1$ и $\hat{\theta}_2$ се определя чрез обратното отношение на техните дисперсии $\sigma_{\hat{\theta}_1}^2$ и $\sigma_{\hat{\theta}_2}^2$:

$$Ef\left(\frac{\hat{\theta}_1}{\hat{\theta}_2}\right) = \frac{\sigma_{\hat{\theta}_1}^2}{\sigma_{\hat{\theta}_2}^2} \quad (10)$$

Смисълът на свойството достатъчност, въведено както и ефективността от Р.Фишер [9] се свежда до оценка на количеството информация, която се съдържа в статистическата извадка. Достатъчна статистика $\hat{\theta}_D$ се нарича тази, която съдържа в себе си цялата информация за θ . Въпросите, свързани с достатъчните статистики изискват по-солидна математическа подготовка [7, 8, 9]. Според тези източници много от известните статистики, включително \bar{x} и S^2 представляват достатъчни статистики.

На базата на така направените разсъждения и доказателства са определени условията за изпълнение на общите свойства на оценките, което е една добра основа за използване на всички статистики, посочени в дясната страна на Табл.1 в следващите изследвания.

IV.Обект и задачи на изследването

4.1. Обща постановка

Изясняването на проблема за определяне на екологичните характеристики на морските транспортни дейности предполага решаването на редица конкретни експлоатационни задачи за оценка на влиянието на тези дейности върху екологичните състояния на атмосферата, хидросферата и крайбрежната суша. В разгледания от нас проблем изхождаме от определението за пристанищна територия в ЗМПВВПНРБ: „Това е територията ограничена от сухопътната граница на кейовия фронт, върху чиято територия са разположени всички пристанищни елементи като оперативна зона, открити и закрити складове, подкранови пътища, железопътни линии и автомобилни пътища, пристанищни съоръжения и др.“ [19]

В конкретния случай ние разглеждаме разтоварния процес на фосфоритна скала (фосфорит) в пристанищния терминал „Варна-запад“ от кораб през гумено-лентов транспортър до цех за обработка на товара „сушене-мелене“ на завода за производство на минерални торове „Агрополихим“, който е подробно описан в [3].

Разтоварването на фосфатната скала (фосфорит) от кораби е процес, който често представлява предизвикателство за индустриалните оператори и регулаторите на околната среда. С растящата глобализация, търсенето на фосфорни продукти за производство на торове и хранителни добавки непрекъснато расте, а заедно с това се увеличава и

интензивността на морските превози на фосфорит. Тази интензивност не може да не се отрази на произшествията и аварията на море и по суша (това вече е доказан факт) [17], което изисква решаването на задачите за контрола и предотвратяването на възможните негативни въздействия върху околната среда и обслужващия персонал.

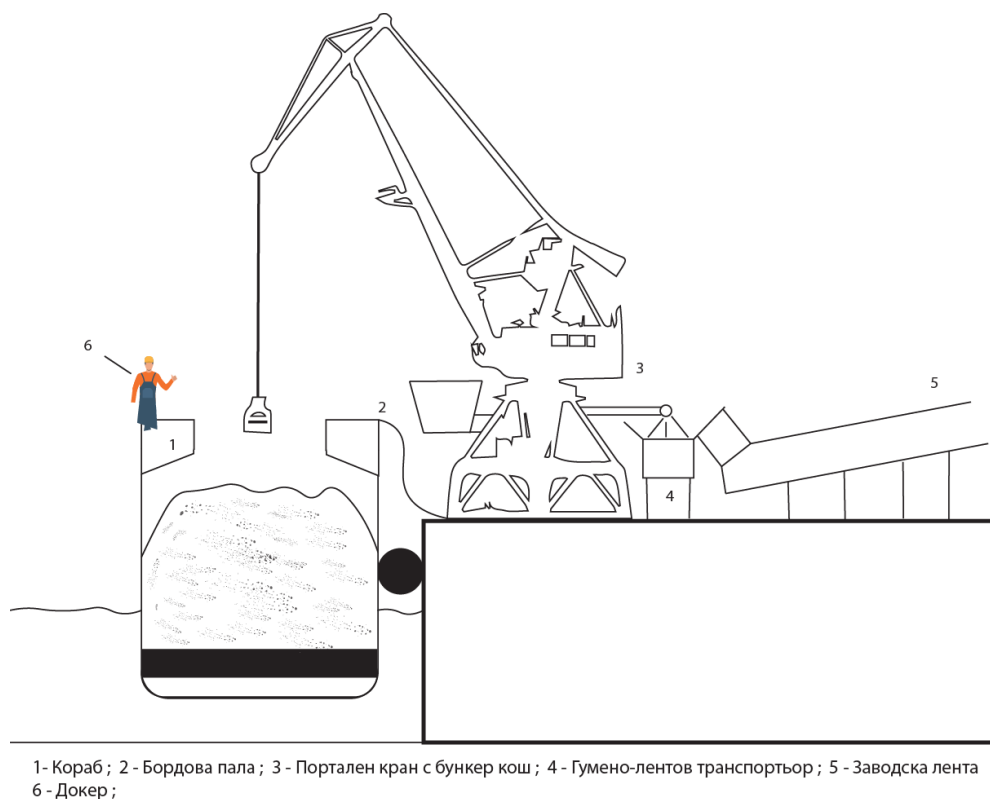
4.2. Описание на технологичния процес, работните места и функциите на лицата по контрол и изпълнение на операциите при разтоварването на фосфорит от кораб за насипни товари на пристанищен терминал „Варна-запад“.

4.2.1. Технологичен процес

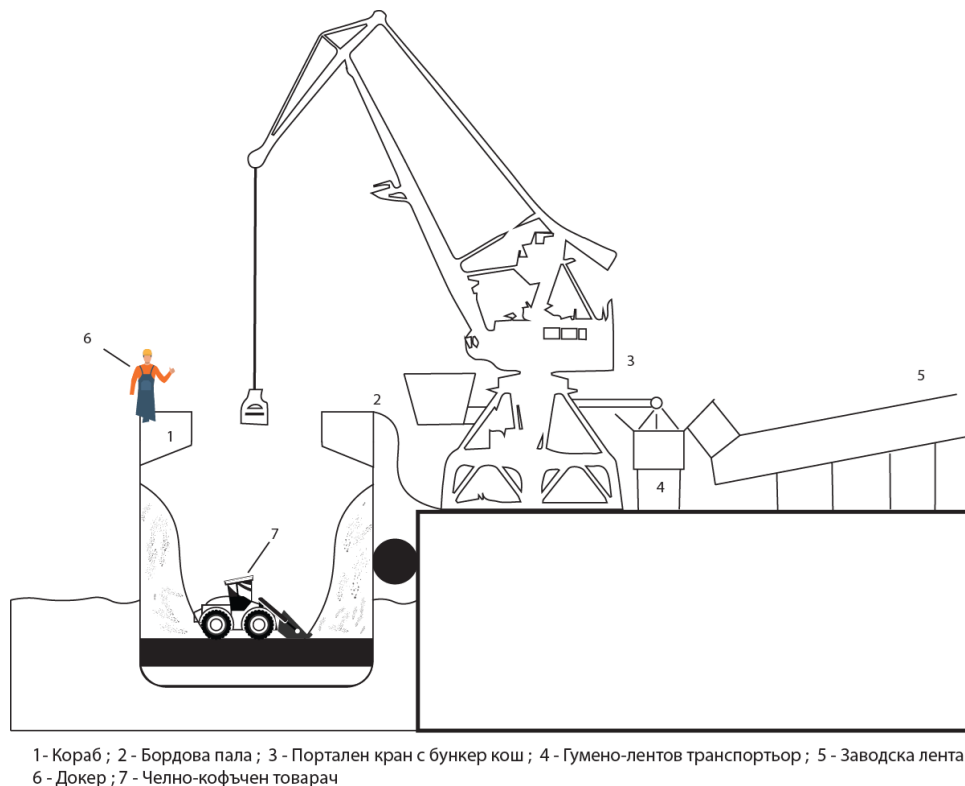
Технологичният процес на разтоварване се осъществява на три слоя, в зависимост от нивата на насипния товар (фосфорит) в трюма на кораба на пристанищния кей.

Операция 1. Разтоварване на ниво (слой 1)

Чрез кранов грайфер фосфоритът се разтоварва от хамбара на кораба. Кранистът изсипва съдържанието на грайфера в бункера на конвейерната установка. От там рудата се подава на гумена транспортна лента. Лентата пренася суровината до цех „Сушене и мелене“ на завода за производство на минерални торове. На фиг.1 е показан товаро-разтоварния процес при работа на I слой.



Фиг. 1. Товаро-разтоварен процес на ниво първо (слой I), при разтоварване на фосфорит от кораб



Фиг. 2. Товаро-разтоварен процес на II слой при разтоварване на фосфорит от кораб

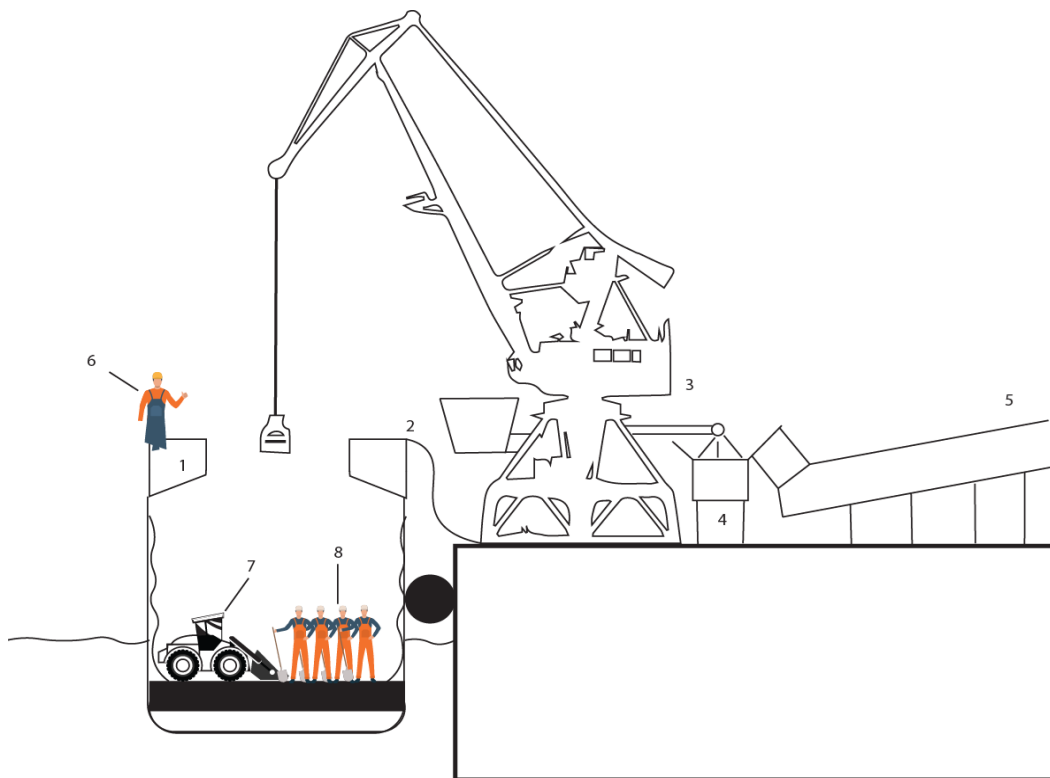
Операция 2 – разтоварване на II слой

Следвайки указанията на докер (сигналчик), водачът на кофъчния товарач избутва стоката от периферията към центъра на хамбара. Чрез кранов грайфер фосфоритът се разтоварва от хамбара на кораба. Кранистът изсипва съдържанието на грайфера в бункера на конвейерната установка. От там рудата се подава на гумена транспортна лента. Лентата пренася суровината до цех „Сушене и мелене“ на завод за производство на минерални торове.

На фиг. 2 е показан товаро-разтоварния процес при работа на II слой.

Операция 3 – разтоварване на III слой

Остатъците от товар, които не могат да бъдат преместени от кофъчния товарач в просвета на люка се събират ръчно с лопати, метли и кофи от докерите(обирачи). Следвайки указанията на докер (сигналчик), водачът на кофъчния товарач избутва стоката от периферията към центъра на хамбара. Чрез кранов грайфер фосфоритът се разтоварва от хамбара на кораба. Кранистът изсипва съдържанието на грайфера в бункера на конвейерната установка. От там рудата се подава на гумена транспортна лента. Лентата пренася суровината до цех „Сушене и мелене“ на завода за производство на минерални торове. На фиг. 3 е показан товаро-разтоварният процес при работа на III слой.



1- Кораб ; 2 - Бордова пала ; 3 - Портален кран с бункер кош ; 4 - Гумено-лентов транспортьор ; 5 - Заводска лента
6 - Докер ; 7 - Челно-кофъчен товарач ; 8 - Докери (обирачи)

Фиг. 3. Товаро-разтоварен процес на III слой при разтоварване на фосфорит от кораб

4.2.2 Описание на работните места и функциите на изпълнителите, на които е необходимо провеждането на екологичен контрол.

1) Докер – сигналчик на палубата при осъществяването на операциите 1,2,3. Контролира отварянето на хамбара (трюма на кораба). Оценява работните условия – 0,5 часа. Подава сигнали (указания) на краниста при разтоварване на насипни товари (фосфорит), чрез кранов грайфер в продължение на 6 часа при работно време от 8 часа. Лични предпазни средства, респираторна маска, ръкавици, очила. Естествена вентилация. Гранични стойности на концентрацията на прах от фосфорит и апатит във въздуха в зависимост от степента на замърсяване: 3-5 – 6,0 mg/m³.

2) Кранист – прави проверка, настройка и управление на конвейера. Изчаква в кабината на крана отварянето на трюма на кораба. Оценка на работните условия. В продължение на 0,5 часа проверява състоянието на работното оборудване. Лични предпазни средства, респираторна маска, ръкавици, очила. Естествена вентилация. Проверка, настройване и управление на конвейера – 4 часа от 8 часово работно време. Гранични стойности на концентрацията от фосфорит и апатит във въздуха в зависимост от степента на замърсяване: 3,5 – 6,0 mg/m³

3) Кранист в процес на разтоварване на нива (слоеве) 1,2,3.

Следвайки указанията на сигналчика, разтоварва минералната руда (фосфорита) от кораба чрез кранов грайфер - 4 часа. Изсипва съдържанието на грайфера в бункера на конвейерната установка, откъдето рудата се подава на гумена транспортна лента, след което се пренася до цех „сушене и мелене“ на завода. Лични предпазни средства; респираторна маска, ръкавици, очила. Естествена вентилация. Силно променящи се условия, при които се наблюдават нива на концентрация на прах от 2,5; 13,7 и 15,6 mg/m³. Гранични стойности на концентрацията на фосфорит и апатит във въздуха в зависимост от степента на замърсяването: 3,5 – 6,0 mg/m³

4) Обходчик – Следи и осигурява правилното функциониране на пристанищната транспортна лента от конвейера до връзката и пресипката с транспортната лента на завода за производство на минерални торове „Агрополихим“. Лични предпазни средства, респираторна маска, ръкавици, очила. Естествена вентилация. Продължителност на работа до конвейера 5 часа от 12 часово работно време. Силно променящи се работни условия-концентрация на прах от 2,8 до 18 mg/m³, гранични стойност на концентрацията при конвейера 4,0 – 6,0 mg/m³.

5) Обходчик - Следи и осигурява работата на връзката между пристанищната транспортна лента и транспортната лента на завода (пресипка). Работа до пресипката 6,5 часа. Лични предпазни средства, респираторна маска, ръкавици, очила. Естествена вентилация. Силно променящи се условия – концентрации на прах от 1,8 до 9,7 mg/m³. Гранични стойности на концентрациите на фосфорит и апатит във въздуха в зависимост от степента на замърсяване: 3,5 – 6,0 mg/m³.

6) Преди разтоварването лицата от новата смяна (сигналчик, кранист, обходчик, водач на челен товарач и докери) изчакват в битовата стая до започване на работните си места – 4 часа от работно време 12 часа. Лични предпазни средства, респираторна маска, ръкавици и очила. Естествена вентилация. Относително стабилни нива на концентрацията на прах : от 1,0 до 6,0 mg/m³. Гранични стойности на концентрацията на фосфорит и апатит във въздуха в зависимост от степента на замърсяване 3,5-6,0mg/m³.

7) Докер в трюма при разтоварване на ниво 3. Подпомага работата на водача на челния товарач при преместването на минералната руда (фосфорит) от периферията към центъра на хамбара на кораба в продължение на 6 часа при работно време 12 часа. Лични предпазни средства, респираторна маска, ръкавици, очила. Естествена вентилация. Силно променящи се условия – концентрации на прах от 2,5; 3,5; 12 mg/m³. Гранични стойности на концентрациите на фосфорит и апатит във въздуха в зависимост от степента на замърсяване: 3,5 – 6,0 mg/m³.

8) Водач на товарач (шофьор) в трюма на кораба при разтоварване на второ и трето ниво. Проверява състоянието на работното оборудване – 0,5 часа. Следвайки указанията на сигналчика, водачът на товарача избутва от периферията към центъра на хамбара минералната руда (фосфорит). Време на работа - 6 часа. Лични предпазни средства, респираторна маска, ръкавици, очила. Естествена вентилация. Силно променящи се

условия – концентрации на прах от 2,5; 3,5; 10; 12 mg/m³. Гранични стойности на концентрациите на фосфорит и апатит във въздуха в зависимост от степента на замърсяване: 3,5 – 6,0 mg/m³.

4.3. Експериментално изследване за определяне на концентрацията на прах във въздуха на работната среда (инхалабилна фракция – прах от фосфорит)

4.3.1. Цел на експерименталните (контролни) изследвания.

Определяне на концентрацията на прах (инхалабилна фракция прах от фосфорит) във въздуха на работната среда на работните места на участниците в процеса на разтоварване на фосфоритна скала (фосфорит) от борда на кораб за насипни товари.

4.3.2. Обект на контрол и описание на състава на контролираните вещества.

Палуба и трюм на кораб за насипни товари „NEW LEGACY” в пристанище Варна-запад на първо кейово място при механично разтоварване на утаечна минерална руда. Изследва се прах от апатит и фосфорит, съдържащ под 2% свободен кристален силициев диоксид в респирабилната фракция. Крайна цел – концентрация на съдържанието на утаечен минерален фосфорит в mg/m³. Оценява се и среднопредетеглената по време концентрация прах и се определят физическите параметри на работната среда по време на вземането на контролните проби.

4.3.3. Технологичен процес

Контролът на работната среда се осъществява по описания в 4.2.1 технологичен процес по време на работа на лицата, участващи в разтоварването при изпълнение на конкретните дейности по технологичните операции (нива 1,2,3)¹.

4.3.4. Технологични средства и схема за измерване

В процеса на контрола са използвани следните технически средства:

1. Дебитомер за въздух вграден в аспиратор „Арех Casella” с касета за вземане на индивидуални или стационарни проби.
2. Цифров барометър „Testo”511.
3. Комбиниран анемометър „KimoAmi 300 STD “
4. Везна електронна Roadwog ХА 60/220.
5. Филтри „Whatman”- тип GF/A 25mm.

За всяко изследвано лице се съставя следният протокол:

Таблица 2. Замервания направени на работното място на сигналчика

Атмосферно налягане, hPa	1031,2
Температура , °C	12,3
Дебит на помпата L, dm ³ /min	2,0
Време за вземане на пробата M, min	120
Време на експозиция, h	6,0
Маса на празен филтър, g ₁ , mg	4854,15
Маса на филтър с прах, g ₂ , mg	4857,72

¹ Контролните измервания са извършени от „Инфостил“ООД - Добрич по заявка на „Пристанище Варна“ ЕАД, която ни предостави резултатите за научен анализ.

Таблица 3. Краен резултат в mg/m^3

Изчислена концентрация на прах	C	17,37
Среднопретеглена по време концентрация	C_{8h}	13,02
Разширена неопределеност	U_0	0,24

4.3.5. Резултати от измерванията

Таблица 4.

Показатели Работни места	Протоколи дата	Концентрация изчислена mg/m^3	Концентр. средна по време	Гранична стойност
Сигналчик	13897 – 01 03.02.2023	2,67	2,01	3,5
Кранист-настройка конвейера	13897 – 01 03.02.2023	2,51	1,26	3,5
Кранист - разтоварване	13897 – 01 03.02.2023	2,41	1,21	3,5
Обходчик при конвейера	13897 – 01 03.02.2023	2,86	1,79	4,0
Обходчик при пресипката	13897 – 01 03.02.2023	1,28	0,53	3,5
Всички в битовата стая	13897 – 01 03.02.2023	1,07	0,53	3,5
Сигналчик	13896 – 01 03.02.2023	17,37	13,02	6,0
Кранист настройва конвейера	13896 – 01 03.02.2023	15,58	7,79	6,0
Кранист разтоварване	13896 – 01 03.02.2023	14,79	7,4	6,0
Обходчик при конвейера	13896 – 01 03.02.2023	17,92	11,2	6,0
Обходчик при пресипване	13896 – 01 03.02.2023	9,71	7,89	3,0-6,0
Всички в битовата стая	13896 – 01 03.02.2023	6,4	3,2	3,0-6,0
Докер – работа в трюма с лопата	13903 – 01 10.02.2023	15,27	11,46	6,0
Водач на челен товарач	13903 – 01 10.02.2023	13,73	10,3	6,0
Докер – работа в трюма с лопата	13904 – 01 10.02.2023	3,25	2,44	3,0-6,0
Водач на челен товарач	13904 – 01 10.02.2023	3,05	2,29	3,0-6,0

Преди да предложим методика за статистическа обработка на експерименталните данни, обръщаме внимание на един предварителен анализ на резултатите от контролните

експериментални изследвания. Числените стойности на концентрациите на състава на респирабилната фракция, получени при сравнително постоянни външни условия (Таблица 4), показват възможности за изследване на три групи от данни, които означаваме като x,y,z.

x – резултатите от измервания, при които концентрацията на инхалабилна фракция прах във въздуха на работната среда (прах от фосфорит), включени основно в протокол 13897/03.02.2023, съответства на изискванията на Наредба N13/Д.В. бр.8/30.01.2004г., изм. Д.В. бр.71/2006г., бр.67/2007г., бр.2/2012г.,бр.46/2016.,бр.73/2018 и бр.5/17.01.2020г. и изм. Д.В. бр.47/04.06.2021г. Измерваните стойности на концентрациите в тази група са в границите между 1 и 3 при пределна стойност $3,5\text{mg}/\text{m}^3$.

y – резултати от измервания, при които концентрацията на инхалабилна фракция прах в Протокол 13903/10.02.2023г. във въздуха на работната среда (прах от фосфорит), включени в Протокол 13896/03.02.2023 не съответства на изискванията на Наредба N13/Д.В. бр.8/30.01.2004г. Измерваните стойности на концентрациите в тази група са в границите между 13 и 18 при пределна стойност $6,0\text{mg}/\text{m}^3$.

z – Концентрацията на инхалабилна фракция прах във въздуха на работната среда (прах от фосфорит) е със стойности в зоната на неопределеност между $3,0$ и $9,0\text{mg}/\text{m}^3$ при гранични стойности, посочени като $3,5\text{mg}/\text{m}^3$ или $6,0\text{mg}/\text{m}^3$ – Протокол 13903/10.02.2023г. и Протокол 13904/10.02.2023г.

Тези разсъждения ни навеждат към задачата от техническата диагностика [10, 12, 13], при която се разглеждат възможностите за разпознаване на данни в три класа: добър, недобър и междинен с тенденция на влошаване. За решаването на задачите за класификация на данните по два или повече класа на състояние е уместно да се обърнем към методите за проверка на статистически хипотези [4,12].

Връщайки се към първоначално поставената задача за разработване на модел за анализ на екологично опасните показатели при товаро-разтоварни операции на насипни товари, обръщаме внимание на следните дейности:

1.Разработване на математичен модел на технологичните и контролни дейности при наличие на два или повече класа на постепенно влошаваща се околна среда.

2.Определяне на статистически обосновани граници между последователно влошаващите се класове на състоянието с помощта на методите за статистическо оценяване и проверка на статистически хипотези [4].

V. Статистическо оценяване и статистически проверки на хипотези при изследване на различни статистически групи от данни.

Преди всичко отбелязваме, че определянето на числените характеристики на разпределенията е тривиално и то се свежда във всички случаи до първите начални

моменти (математическо очакване $\mu = \bar{x}$) и вторите централни моменти (дисперсии s^2) и те се пресмятат по известни формули (Табл.1).

Задачите, които предстоят да бъдат решавани при анализ на данни от различни статистики (групи от измервани величини) в конкретно поставената от нас задача, както и във всички задачи при статистически контрол на качеството, свързани с измерването на някаква характеристика (показател) на изследван обект или изделие са:

1) Да се провери дали (и до колко) средната стойност на тази характеристика се различава съществено от конкретни стандартни стойности. От статистическа гледна точка тази задача е известна като „Проверка на хипотези за равенство (или различие) на математическото очакване на случайната величина μ на зададена стойност μ_0 “.

2) Да се провери дали (и до колко) две или повече от анализирани статистически групи принадлежат към една генерална съвкупност или са съществено различни. Тази задача е известна в статистиката като „Проверка на хипотезата за равенство на две (или повече) математически очаквания“.

3) Да се провери разсейването на стойностите на случайните измервания (изразено математически чрез дисперсията s^2) около нейното математическо очакване. Такава задача възниква при оценка на точността на избрания показател или метод на изследване на технологичния процес, а в статистиката тя е известна като „Проверка на хипотезата за равенство на дисперсията на случайна величина на зададена стойност или на дисперсията на друга случайна величина“.

Преди да представим процедурите и алгоритмите за приемане на управляващи решения въз основа на данни от случайни извадки, каквато е поставената пред нас задача, отбелязваме, че приемането на всяко решение се свежда до проверката на едни или други твърдения или предположения за свойствата на случайните величини по резултатите от частични наблюдения. Тези твърдения или предположения се наричат статистически хипотези.

Исторически първата известна статистическа проверка на хипотеза е изследването на Дж.Арбутнот от 1710г., описано в труда му: „Доводи в полза на божествените пророчества, изведени на основата на постоянни и систематични наблюдения над ражданията на двата пола“. Като изследва старателно демографските записи в течение на 82 години, той е показал, че броят на ражданите момчета систематично превишава броят на ражданите момичета. По този начин Арбутнот опровергава хипотезата, считана дотогава за безспорна, че ражданията на момчетата и момичетата са равновероятни, т.е. имат вероятност 0,5. Като доказва статистически безупречно своята хипотеза, Арбутнот се опитва да я обоснове с божествената намеса от гледна точка на божествения закон за моногамното човешко общество. Този първи пример на статистическа проверка на хипотеза, макар и да изглежда куриозен, показва едновременно, че математическата статистика може да обоснове един или друг извод по резултатите от наблюденията, но не може и няма задача да обясни неговата гносеологическа същност [4].

Обикновено статистическата проверка на хипотези се означава с H_1 : (твърдение, че ...).

Така например хипотезата, че математическото очакване е равно на μ_0 се означава с $H: \mu = \mu_0$, а хипотезата, че дисперсиите σ_1^2 и σ_2^2 са равни - $H: \sigma_1^2 = \sigma_2^2$.

След формулирането на хипотезата трябва да се провери вероятността на изразеното твърдение за генералната съвкупност въз основа на данните от частичната съвкупност. Методът, който се използва в теорията на статистическите решения, за да се вземе решение за справедливостта на статистическата хипотеза се нарича „проверка на хипотези“.

Ако дадена хипотеза (например $H: \mu = \mu_0$) не е вярна, трябва да съществува друга противоположна, алтернативна хипотеза, която да е вярна (например $H: \mu > \mu_0$). Обикновено основната хипотеза, която се проверява се означава с H_0 : и се нарича „нулева хипотеза“, а алтернативната хипотеза се означава с H_1 :

В следващия текст ще бъдат представени алгоритмите за решаване на трите поставени задачи за проверка на хипотези, като се следват идеите, изразени в [4, 5, 12].

5.1.Алгоритъм за проверка на хипотезите за равенство (алтернативна-неравенство) на математическото очакване на случайните измервания на зададена (гранична) стойност на измервания показател.

Тъй като в процедурата за контрол на концентрацията на прах от фосфорит във въздуха (Табл.4) са заложили две гранични стойности, е желателно на базата на ограничения брой опити да се проверят хипотезите, че средната стойност на статистическата извадка е достатъчно отдалечена (или близка) до всяка от тези контролни стойности.

Решаваме задачата по следната схема.

5.1.1. Данни

Данните се състоят от n наблюдения на случайната величина (в случая концентрация) $X: x_1; x_2, \dots, x_n$. Броят на наблюденията n може да бъде произволен, но може да бъде избран и така, че да осигурява зададена точност на изследването. Средноаритметичното значение на случайната величина е $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$, а граничната зададена стойност е μ_0 . Задачата се състои в това, да се провери, дали $(\mu_0 - \bar{x}) \leq \delta$, където δ е предварително зададена стойност. В статистическия контрол на качеството на продукцията се поставя въпросът, какъв трябва да бъде минималният обем на извадката, за да се осигури правилно решение с вероятност $P=1 - \alpha$, където α се нарича ниво на значимост, а в теорията на приемане на решения е известна като грешка от първи род. В нашия случай ние решаваме обратната задача – да се определи минималната стойност на грешката (или максималната вероятност) при малък обем на статистиката n .

И двата въпроса могат да се решават по един и същ алгоритъм, който има две разновидности според това, дали е известна дисперсията на генералната съвкупност σ^2 или е пресметната оценката s^2 по малък обем на статистиката.

В първия случай при известна σ^2 , величината:

$$U = \frac{\delta\sqrt{n}}{\sigma} \quad (11)$$

има нормирано нормално разпределение и за нея може да бъде построена двустранна критична област с ниво на значимост α [4, 5]. Ако границите на тази критична област се означат с $-U_{кр}$ и $U_{кр}$, то с вероятност $(1 - \alpha)$ е изпълнено неравенството:

$$-U_{кр} \leq \frac{\delta\sqrt{n}}{\sigma} \leq U_{кр} \quad (12)$$

Величината $U_{кр}$ е такава, че стойността на функцията на стандартното нормално разпределение $\Phi(U_{кр})$, пресметната по интеграла на Лаплас и зададена таблично ([4], табл.2) да бъде равна на $(1 - \alpha)$. Тогава от условието (12) е ясно, че минималният обем на извадката е:

$$n = \frac{U_{кр}^2 \cdot \sigma^2}{\delta^2}, \quad (13)$$

където $U_{кр}$ се избира така, че $\Phi(U_{кр}) = 1 - \alpha$.

Във втория случай при неизвестна σ^2 , а известна нейната оценка s^2 , пресметната по ограничен брой данни, се използва разпределението на Стюдънт. Тогава

$$n = \frac{t^2\left(v, \frac{\alpha}{2}\right) s^2}{\delta^2}, \quad (14)$$

където с $t\left(v, \frac{\alpha}{2}\right)$ е означен пресметнатият и таблично представен квантил на разпределението на Стюдънт при ниво на значимост $\frac{\alpha}{2}$ и брой на степените на свобода $v = n - 1$ ([4], табл.4).

5.1.2. Предположения, използвани при формулирането на критериите и хипотезите.

Правят се следните предположения:

5.1.2.1. X е случайна величина, чието генерално разпределение се подчинява на нормален закон.

5.1.2.2. Наблюденията са независими.

5.1.2.3. Всички наблюдения имат една и съща дисперсия σ^2 , която може да бъде известна или неизвестна.

5.1.2.4. Поради ограничения брой на експерименталните данни не може да бъде потвърдено предположението за нормално разпределение с дисперсия σ^2 . По тази причина задачата се решава при предположение, че дисперсиите са неизвестни, а техните оценки s^2 се пресмятат по наличните малки обеми на данните, които се подчиняват на разпределението на Стюдънт.

5.1.2.5. При проверка на хипотезата за различие на математическото очакване на данните от граничните стойности ($\mu \neq \mu_0$) се използват двустранни доверителни интервали $t\left(v, \frac{\alpha}{2}\right)$, които са квантили на разпределението на Стюдънт при ниво на значимост $\frac{\alpha}{2}$ и брой на степените на свобода $v = n - 1$ ([4], табл.4)

5.1.2.6. При проверка на хипотезите $\mu < \mu_0$ и $\mu > \mu_0$ се използват едностранни доверителни интервали $t(v, \alpha)$.

Поради двустранната симетрия на нулевата и алтернативната хипотеза е достатъчно да се провери само нулевата хипотеза и ако тя не се потвърждава, да се приеме алтернативната.

При формулирането на хипотезите с μ се означава математическото очакване на случайната величина X , а с μ_0 - зададената граница, с която се сравнява средната стойност. Използват се следните хипотези:

- а) двустранен критерий: нулевата хипотеза е $H_0: \mu = \mu_0$, а алтернативната е $H_1: \mu \neq \mu_0$ (фиг.4а);
- б) едностранен критерий: нулевата хипотеза е $H_0: \mu \leq \mu_0$, а алтернативната е $H_1: \mu > \mu_0$ (фиг.4б);
- в) едностранен критерий: нулевата хипотеза е $H_0: \mu \geq \mu_0$, а алтернативната е $H_1: \mu < \mu_0$. (фиг.4в)



Фиг. 4. Нива на значимост и доверителни интервали при двустранен и едностранен критерий

5.1.3. Процедура за проверка на хипотезите за равенство на математическото очакване на гранична стойност при неизвестни дисперсии на генералната съвкупност

5.1.3.1. Пресмятат се оценките на математическото очакване и дисперсията:

$$\left. \begin{aligned} \bar{x} &= \frac{1}{n} \sum_{i=1}^n x_i \rightarrow \mu \\ s^2 &= \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \rightarrow \sigma^2 \end{aligned} \right\} \quad (15)$$

$$T = \frac{(\bar{x} - \mu_0) \sqrt{n}}{s} \quad - \text{нормирана ширина на статистическия интервал.}$$

Величината T има разпределение на Стюдънт с $v = n - 1$ степени на свобода.

5.1.3.2. От таблиците на разпределение на Стюдънт (напр. табл.4 от Приложение на [4]) се определя квантилът на разпределението на Стюдънт $t\left(v, \frac{\alpha}{2}\right)$ за двустранен критерий или $t(v, \alpha)$ за едностранен критерий. Двустранен критерий се използва, ако се проверява условието за това, че величината μ_0 е извън двустранния доверителен интервал (фиг.4а). Едностранни критерии се използват за проверка на условията $\mu \leq \mu_0$ (фиг.4б) или $\mu \geq \mu_0$ (фиг.4в).

5.1.3.3. Проверката на хипотезите се изпълнява по следните правила:

а) двустранен критерий при нулева хипотеза $H_0: \mu = \mu_0$ и алтернативна хипотеза $H_1: \mu \neq \mu_0$. Нулевата хипотеза се приема (не се отхвърля), ако $|T| \leq t\left(v, \frac{\alpha}{2}\right)$. Алтернативната хипотеза се приема, ако $|T| > t\left(v, \frac{\alpha}{2}\right)$.

б) едностранен критерий при нулева хипотеза $H_0: \mu \leq \mu_0$ и алтернативна хипотеза $H_1: \mu > \mu_0$. Нулевата хипотеза се приема, ако $T \leq t(v, \alpha)$. Нулевата хипотеза се отхвърля, а се приема алтернативната, ако $T > t(v, \alpha)$.

в) едностранен критерий при нулева хипотеза $H_0: \mu \geq \mu_0$ и алтернативна хипотеза $H_1: \mu < \mu_0$. Нулевата хипотеза се приема, ако $T \geq -t(v, \alpha)$. Нулевата хипотеза се отхвърля, а се приема алтернативната, ако $T < -t(v, \alpha)$.

Процедурите за проверка на хипотези при двустранни и едностранни критерии са относително устойчиви към нарушение на предпоставката за нормалност на данните. Причината за това е, че t - разпределението на Стюдънт асимптотично се стреми към нормалното, когато $n \rightarrow \infty$. При малък брой на данните, използването на това разпределение е единственият начин за проверка на хипотези, а с увеличаване на обема на данните величината T е асимптотично нормална, което е предпоставка за по-точни решения, основаващи се на уравнение (13).

5.2. Алгоритъм за проверка на хипотези за равенство (различие) на две или повече математически очаквания

5.2.1. 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. Тази задача се решава по процедурите а) и б) за всички възможни комбинации.

5.2.2. Данни

Сравняват се две или повече извадки, взети съответно от генералните съвкупности X, Y, Z (при три предполагаеми класа) с обеми n_1, n_2, n_3 . Независимо от броя на класовете, проверката на хипотезите за равенство на математическите очаквания се изпълнява за всички възможни комбинации от два класа.

5.2.3. Предположения при формулиране на критериите и хипотези за проверка

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

2) Наблюденията от сравняваните извадки са независими.

3) Дисперсиите $\sigma_1^2, \sigma_2^2, \sigma_3^2, \dots$ на извадките могат да бъдат известни или да се определят експериментално техните оценки $s_1^2, s_2^2, s_3^2, \dots$. 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 σ_i^2 . Отбелязваме, че ако сравняваните извадки имат голям обем ($n_i > 30$), разпределението на средните аритметични $\bar{x}, \bar{y}, \bar{z}$ се стреми към нормалното, независимо от разпределението на случайните величини x, y, z , което се обяснява с централната гранична теорема [4, 5]. 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 обозначи 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, например концентрации на прах, съответстващи на Наредба 13 (X) или несъответстващи на изискванията на Наредба 13 (Y).

5.2.4. Процедура (алгоритъм) за проверка на хипотезите за равенство на математическите очаквания при малки обеми на извадките с неизвестни дисперсии

5.2.4.1. Предполага се, че дисперсиите са равни $\sigma_1^2 = \sigma_2^2$.

Този случай се среща, когато двете извадки са получени при едни и същи условия (температура, налягане, влажност и др.). 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

$$\bar{x} = \frac{1}{n_1} \sum_{i=1}^{n_1} x_i ; \quad \bar{y} = \frac{1}{n_2} \sum_{i=1}^{n_2} y_i \quad (16)$$

и дисперсиите:

$$s_1^2 = \frac{1}{n_1-1} \sum_{i=1}^{n_1} (x_i - \bar{x})^2 ; \quad s_2^2 = \frac{1}{n_2-1} \sum_{i=1}^{n_2} (y_i - \bar{y})^2 \quad (17)$$

Пресмята се нормираното разстояние между разпределенията:

$$T = \frac{(\bar{x} - \bar{y})}{\sqrt{(n_1-1)s_1^2 + (n_2-1)s_2^2}} \sqrt{\frac{n_1 \cdot n_2 (n_1 + n_2 - 2)}{n_1 + n_2}} \quad (18)$$

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

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

Проверка на хипотезите се извършва по следните правила:

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

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 $T < t(v, \alpha)$. Ако $T \geq t(v, \alpha)$, 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 $T > -t(v, \alpha)$. В противен случай, ако $T \leq -t(v, \alpha)$ нулевата хипотеза се отхвърля и се приема алтернативната.

5.2.4.2. Дисперсиите σ_1^2 и σ_2^2 са неизвестни и не са равни ($\sigma_1^2 \neq \sigma_2^2$).

1) Изчисляват се \bar{x} , \bar{y} и s_1^2 , s_2^2 по формулите (16) и (17).

2) Изчислява се величината:

$$T = \frac{(\bar{x} - \bar{y})}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad (19)$$

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

$$v = \frac{(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2})^2}{\frac{(\frac{s_1^2}{n_1})^2}{n_1+1} + \frac{(\frac{s_2^2}{n_2})^2}{n_2+1}} - 2 \quad (20)$$

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

3) Проверката на хипотезите се извършва по същите правила, както в т.5.2.3 по стойности на T и v , изчислени по формули (19) и (20).

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

5.3.1 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). В конкретно провежданите от нас изследвания тази задача, заедно с разгледаните в (5.1) и (5.2) методи може да бъде основа за преразглеждане на граничните стойности между класовете на състояние, когато те са повече от два.

5.3.2. 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, \dots, x_n , а извадката от генералната съвкупност Y е с n_2 наблюдения, $\{Y\}$: y_1, y_2, \dots, y_n ,

5.3.3. Предположения и хипотези

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

Ако със σ_1^2 се обозначи 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: $H_0: \sigma_1^2 \geq \sigma_2^2$; $H_1: \sigma_1^2 < \sigma_2^2$.

5.3.4. 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. [9, 33]

Ако със s_1^2 се обозначи оценката на дисперсията σ_1^2 , а със s_2^2 - оценката на дисперсията σ_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:

$$\frac{1}{F\left(\frac{\alpha}{2}, \nu_1, \nu_2\right)} < \frac{s_1^2}{s_2^2} < F\left(\frac{\alpha}{2}, \nu_1, \nu_2\right), \quad (21)$$

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 ([4]Табл.5, Табл.6). 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 (21) винаги се изпълнява, поради което не е нужно да се проверява. Това означава, че трите случая се свеждат до проверка на една хипотеза:

$$H_0: \sigma_1^2 = \sigma_2^2 \text{ и } H_1: \sigma_1^2 > \sigma_2^2; \quad (22)$$

5.3.5. Алгоритъм за проверка на хипотези

1) Изчисляват се оценките:

- на математическите очаквания

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

- на дисперсиите

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

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).

3) Изчислява се броят на степените на свобода за числителя $\nu_1 = n_1 - 1$ и за знаменателя $\nu_2 = n_2 - 1$. При зададено ниво на значимост α от таблиците на разпределение на Фишер (табл.5 или 6 в Приложение на [11]) се отчита $F(\alpha, \nu_1, \nu_2)$.

4) 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 $H_1: \sigma_1^2 > \sigma_2^2$.

VI.Реализация на алгоритмите за проверка на хипотези за контрол на концентрацията на прах от фосфорит във въздуха при товаро-разтоварни дейности

6.1. Проверка на хипотезите за равенство или различие на математическите очаквания на изследваните групи от данни и граничните контролни стойности

Тази част от изследването, осъществена по алгоритъма за проверка на хипотези (5.1) има за цел да се провери взаимното разположение на центровете на разпределенията на експерименталните данни μ_x , μ_y , μ_z и априорно зададените гранични стойности на концентрацията на прах $\mu_0 = 3,5$ и $\mu_0 = 6$ и да се потърсят статистически обосновани граници между класовете при повече от два класа.

6.1.1. Данни от клас X. Резултати от измерванията с концентрации на прах от фосфорит във въздуха, съответстваща на изискванията на Наредба 13/ДВ бр.8/30.01.2004г.

Част I. Проверка на хипотезите при зададено ниво на значимост α

Клас x, $\mu = \bar{x} = 2,133$; $n = 6$; $s^2 = 0,579$; $s = 0,76$

1.1. Нулева хипотеза $H_0: \mu = \mu_0$; алтернативна хипотеза $H_1: \mu_1 \neq \mu_2$.

Проверка по двустранен критерий при ниво на значимост $\alpha = 0,1$; ($P = 0,9$), при двустранен критерий $\frac{\alpha}{2} = 0,05$; ($P = 0,95$)

а) гранична стойност $\mu_0 = 3,5$

$$T = \frac{(\mu - \mu_0)\sqrt{n}}{s} = \frac{(2,133 - 3,5) \cdot 2,45}{0,76} = -4,41; |T| = 4,41$$

От табл.4 в [4] определяме при $\nu = n - 1 = 5$ и $\frac{\alpha}{2} = 0,05 \Rightarrow t(5; 0,05) = 2,571$

Проверка: Ако $|T| < t(5; 0,05)$ нулевата хипотеза $H_0: \mu = \mu_0$ се приема.

Ако $|T| \geq t(5; 0,05)$ приема се алтернативната хипотеза $H_1: \mu \neq \mu_0$.

В нашия случай $|T| = 4,41 > t(5; 0,05) = 2,571$. Приема се алтернативната хипотеза $H_1: \mu \neq \mu_0$ с вероятност $P = 0,9$

б) гранична стойност $\mu_0 = 6$; $T = \frac{(2,133 - 6) \cdot 2,45}{0,76} = -12,47$

Проверка на нулевата хипотеза: $|T| = 12,47 > t(5; 0,05) = 2,571$, приема се алтернативната хипотеза $H_1: \mu \neq \mu_0$ с вероятност $P = 0,95$

Нулева хипотеза $H_0: \mu < \mu_0$. Работи се по едностранен критерий с ниво на значимост $\alpha = 0,1$

а) $\mu_0 = 3,5$; $T = -4,41$; $\alpha = 0,1$

От табл.4 в [2] определяме $t(5; 0,1) = 2,015$.

Проверка на нулевата хипотеза $H_0: \mu < 3,5$. Тъй като $T = -4,41 < t(5; 0,1) = 2,015$ нулевата хипотеза се приема с $P = 0,9$.

б) $\mu_0 = 6$; $T = -12,47$; $\alpha = 0,1$

От табл.4 в [11] определяме $t(5; 0,1) = 2,015$.

Проверка на нулевата хипотеза

$T = -12,47 < t(5; 0,1) = 2,015 \Rightarrow$ приема се нулевата хипотеза $H_0: \mu < 6$ с $P = 0,9$.

Част II. Определяне на границите на валидност на приетата нулева или алтернативна хипотеза

При приемането на конкретни решения в задачите за контрол на състоянието на изследвания обект (съоръжение) по измерените стойности на един или няколко показателя (признаци) е важно познаването на границите на валидност на избраната нулева или алтернативна хипотеза, т.е. става дума за това с какъв запас разполага страната (лицето), приемаща решението. Като имаме предвид случайния характер на изменение на векторите на наблюдение е добре да търсим някаква „вероятна“ мярка на разстоянието в координатното пространство на признаците. В системната теория при многомерен вектор на наблюдение (при два или повече признака) като вероятностна мярка на разстоянието между различните състояния се използва т.н. „квадратично разстояние по Махаланобис“ [10, 16, 30]. При едномерни сигнали (напр. концентрация на примеси във въздуха), какъвто е нашият случай, понятието „разстояние между класовете“ може да се измерва по стойностите на нивата на значимост α или на съответстващите им вероятности P . Това дава възможност за използване на части от идеите и процедурите на статистическата проверка на хипотези за определяне на вероятностните граници на валидност на приетата хипотеза. Припомняме, че в статистическия контрол на качеството [4] се търси решение на задачата за определяне на минималния обем на извадката от данни n при известни оценки на дисперсията s^2 и разликата между контролната и средна стойност $\delta = \mu_0 - \bar{x}$ (формули 13 и 14).

Тук ние си поставяме обратната задача: по известни параметри на разпределението ($\mu_0 = \bar{x}$, n , s^2 и пресметнати стойности на доверителния интервал T) да се определи минимално допустимата (критична) стойност на нивото на значимост $\alpha_{кр}$ или максималната стойност на вероятността $P_{max} = 1 - \alpha_{кр}$. За решаването на тази задача се използват табличните пределни стойности на квантилите на разпределението на Стюдънт, за които са изпълнени условията (Табл.4 [4]):

- при проверка на двустранен критерий:

$$H_1: \mu \neq \mu_0, \text{ ако } |T| > t\left(v, \frac{\alpha_{кр}}{2}\right) \quad (23)$$

- при проверка на едностранен критерий:

$$H_0: \mu < \mu_0, \text{ ако } |T| < t(v, \alpha_{кр}) \quad (24)$$

Техниката за определяне на вероятностните мерки $\alpha_{кр}$ или P_{max} се свежда до търсенето на табличните стойности на квантилите $t(v, \alpha_{кр})$, при които са допустими горните неравенства.

Резултати по част II

2.1. Граница на валидност на хипотезата за различие на математическото очакване на измерените стойности μ от граничните (допустими) стойности на концентрацията на прах във въздуха μ_0 при алтернативна хипотеза $H_1: \mu \neq \mu_0$ и двустранен критерий.

а) при гранична стойност $\mu_0 = 3,5$

$$|T| = 4,41 \leq t(5; 0,01/2) = 4,43$$

Алтернативната хипотеза $\mu \neq \mu_0$ е валидна до нива на значимост $\frac{\alpha}{2} = 0,005$ или $P=0,995$. При $\frac{\alpha}{2} < 0,005$ алтернативната хипотеза не се приема (тя е невалидна).

б) при гранична стойност $\mu_0 = 6$

$$|T| = 12,4 \leq t(5; 0,01/2) = 12,4$$

Алтернативната хипотеза е валидна с нива на значимост до $\frac{\alpha}{2} = 0,0005$ с минимална вероятност $P_{\max} = 0,9995$.

2.2. Валидност на нулевата хипотеза $H_0: \mu < \mu_0$ по едностранен критерий:

а) гранична стойност $\mu_0 = 3,5$

$$T = -4,4 \leq t(5; 0,01) \approx 4,43$$

Нулевата хипотеза е валидна до нива на значимост $\alpha = 0,01$, което отговаря на вероятност $P = 0,99$.

б) гранична стойност $\mu_0 = 6$

$$T = -12 \leq t(5; 0,0005) \approx 12$$

Нулевата хипотеза е валидна до нива на значимост $\alpha = 0,0005$, което отговаря на вероятност $P = 0,9995$.

6.1.2. Данни от клас Y. Резултати от измервания с концентрация на прах от фосфорит във въздуха, несъответстваща на изискванията на Наредба 13/ДВ.бр.8/30.01.2004г.

Част I. Проверка на хипотезите при зададено ниво на значимост α

$$\mu = \bar{y} = 15,78; s^2 = 2,518; s = 1,59; n = 6; \alpha = 0,1$$

1.1. Нулева хипотеза $H_0: \mu = \mu_0$

алтернативна хипотеза $H_1: \mu \neq \mu_0$

Проверка при двустранен критерий, при който се работи при нива на значимост $\frac{\alpha}{2} = 0,05; P = 0,95$.

а) гранична стойност $\mu_0 = 3,5$

$$T = \frac{(\mu - \mu_0)\sqrt{n}}{s} = 18,92$$

от табл.4 [10] при $v = 5; \alpha = 0,05; t(5; 0,05) = 2,571$

Проверка $|T| = 18,92 > t(5; 0,05) = 2,571$

Нулевата хипотеза се отхвърля и се приема алтернативната $H_1: \mu \neq \mu_0$ с вероятност $P=0,95$.

б) гранична стойност $\mu_0 = 6$

$$T = 15,06; t(5; 0,05) = 2,571$$

Проверка: $|T| = 15,06 > t(5; 0,05) = 2,571$

Нулевата хипотеза се отхвърля и се приема алтернативната $H_1: \mu \neq \mu_0$ с вероятност $P = 0,95$.

1.2. Проверка на хипотезата $H_0: \mu < \mu_0$ по едностранен критерий с ниво на значимост $\alpha = 0,1$

а) гранична стойност $\mu_0 = 3,5$; $T = 18,92$; $t(5; 0,1) = 2,015$.

Проверка: $T = 18,92 > t(5; 0,1) = 2,015$. Приема се алтернативната хипотеза $H_1: \mu \geq \mu_0$ с вероятност $P = 0,9$.

б) гранична стойност $\mu_0 = 6$; $T = 15,06$; $t(5; 0,1) = 2,015$

Проверка: $T = 15,06 > t(5; 0,1) = 2,015$. Приема се алтернативната хипотеза $H_1: \mu \geq \mu_0$ с вероятност $P = 0,9$.

Част II. Определяне на границите на валидност на приетите нулева и алтернативна хипотези.

2.1. Граница на валидност на хипотезата за различие на математическото очакване на измерените стойности $\mu = \bar{x}$ от пределните (допустими) стойности на концентрацията на прах във въздуха при алтернативна хипотеза $H_1: \mu \neq \mu_0$

а) при гранична стойност $\mu_0 = 3,5$; $|T| = 18,92 \gg t(5; 0,0005)$

б) гранична стойност $\mu_0 = 6$; $|T| = 15,06 \gg t(5; 0,0005)$

Това означава, че приемането на алтернативната хипотеза $H_1: \mu \neq \mu_0$ е с вероятност $P \approx 1$ за целия диапазон на пресмятане и таблично представяне на квантилите на разпределението на Стюдънт.

2.2. Валидност на нулевата хипотеза $H_0: \mu \geq \mu_0$ по едностранен критерий.

а) при гранична стойност $\mu_0 = 3,5$; $T \gg t(5; 0,0005)$

б) гранична стойност $\mu_0 = 6$; $T \gg t(5; 0,0005)$

В този случай приемането на нулевата хипотеза $H_0: \mu \geq \mu_0$ е с вероятност $P \approx 1$ за целия диапазон на пресмятане и таблично представяне на квантилите на разпределението на Стюдънт.

Изводът от проверката на валидността на двете разгледани хипотези за приемане на решения е, че разстоянието между центровете на разпределения на данните от двата класа, условно наречени „добър и недобър“ е толкова голямо, че би могло да се помисли, дали не е целесъобразно да бъде въведен един междинен клас, наречен „допустим, но опасен“. Още повече, че между измерените данни има такива, които дават възможност за формирането на такъв клас.

6.1.3. Предполагаме клас Z – „допустим, но опасен“. Данни, разположени между или близо до граничните стойности на основните класове.

Част I. Проверка на хипотезите при зададено ниво на значимост α

$\mu = \bar{z} = 5,603$; $s^2 = 9,852$; $s = 3,14$; $n = 4$; $v = 3$.

Нулева хипотеза $H_0: \mu = \mu_0$

алтернативна хипотеза $H_1: \mu \neq \mu_0$

Проверка при двустранен критерий, при който се работи при нива на значимост $\frac{\alpha}{2} = 0,05$; ($P=0,95$).

а) гранична стойност $\mu_0 = 3,5$

$$T = 1,34; \quad t\left(v, \frac{\alpha_{\text{кр}}}{2}\right) = t(3; 0,05) = 2,353$$

$$|T| = 1,34 < t(3; 0,05) = 2,353$$

Приема се нулевата хипотеза $H_0: \mu = \mu_0$, тъй като математическото очакване на измерените стойности е много близо до долната гранична стойност $\mu_0 = 3,5$

б) гранична стойност $\mu_0 = 6$

$$T = -0,26; \quad t(3; 0,05) = 2,353$$

$$|T| = 0,26 < t(3; 0,05) = 2,353$$

Приема се нулевата хипотеза $H_0: \mu = \mu_0$, тъй като математическото очакване на групата на измерените стойности е много близо до горната гранична стойност $\mu_0 = 6$.

Средната стойност на измервания (контролиран) показател за междинния предполагаем клас е между долната $\mu = 3,5$ и горната $\mu_0 = 6$ гранични стойности. Поради това и в двата разгледани случая не може да се приеме хипотезата $H_1: \mu \neq \mu_0$. Ако в следващите изследвания се потвърди хипотезата за възможно съществуване на междинен клас, би трябвало граничните стойности да се променят, за което има статистическа възможност.

6.2. Реализация на алгоритмите за проверка на хипотези за равенство (различие) на две или повече математически очаквания

Тази част от изследването, осъществена по алгоритъма за проверка на хипотези (5.2) има за цел проверката на възможностите за формиране на резултатите от проведените експериментални измервания на посочените в (5.1) три класа (X, Y, Z). Без да разглеждаме задачите за проверка на хипотезите за нормалност на разпределенията на всяка от експерименталните групи или за тяхната независимост, ние приемаме, че най-съществените условия за различимост на отделните разпределения се свеждат до проверката на хипотезите за различие на техните центрове, т.е. на математическите им очаквания [4, 9, 10]. Проверките на хипотезите за равенство (различие) на математическите очаквания се извършват за всяка двойка от класовете на състояние.

6.2.1. Проверка на хипотези за различие на математическите очаквания на групите от данни, съответстващи (X) и несъответстващи (Y) на Наредба 13/ДВ. Бр.8/30.01.2004г. (Табл.2 и 3)

1) Числени стойности на оценките на параметрите на разпределенията

$$\text{Клас } X: \mu_x = \mu_1 = \bar{\mu} = 2,133; \quad s_x^2 = 0,579; \quad s_x = 0,76; \quad n_x = 6$$

$$\text{Клас } Y: \mu_y = \mu_2 = \bar{y} = 15,777; \quad s_y^2 = 2,518; \quad s_y = 1,587; \quad n_y = 6$$

2) Хипотези:

а) двустранен критерий: $H_0: \mu_x = \mu_y; \quad H_1: \mu_x \neq \mu_y$

б) едностранен критерий: $H_0: \mu_x \leq \mu_y$; $H_1: \mu_x > \mu_y$

Работи се по процедурата за неизвестни, но равни дисперсии. Подробната реализация на алгоритмите за проверка на хипотезите по едностранен и двустранен критерий е представена на фиг.5,6,7.

табл.1 - I-ви клас - Съответствие на Наредба №13/30.01.2004					
№	Работно място и изпълнител	изчисл. концентр.	средно прет. по	неопределеност	пределна стойност
1	Сигналчик. От палуата дава указания на краниста- 6 ч.	2,67	2,01	0,04	3,50
2	Кранист.Проверка, настройка и управление на конвейра (в трюма). Продължителност 4 часа	2,51	1,26	0,02	3,50
3	Кранист. Разтоварване по указание на сигналчика- 4 ч.	2,41	1,21	0,02	3,50
4	Обходчик-до конвейра. Следи работата на трансп.лента- 5ч	2,86	1,79	0,03	4,00
5	Обходчик при пресипване. Следи работата на лентата - 5 ч	1,28	1,04	0,02	3,50
6	Всички работници изчакват в битовата зала - 4 ч.	1,07	0,53	0,01	3,50
	μ	2,133	1,307	0,023	3,583



табл.2 - II-ри клас - Не съответствие на Наредба №13/30.01.2004					
№	Работно място и изпълнител	изчисл. концентр.	средно прет. по	неопределеност	пределна стойност
1	Сигналчик. От палуата дава указания на краниста при разтоварване 6 часа	17,37	13,02	0,24	6,00
2	Кранист.Проверка, настройка и управление на конвейра (в трюма) - 4 часа	15,58	7,79	0,14	6,00
3	Кранист. Разтоварва насипни товари от трюма по указание на сигналчика - 4 часа.	14,79	7,40	0,13	6,00
4	Обходчик-Осигурява функционирането на трансп.лента (до конвейра) - 5 часа	17,92	11,20	0,19	6,00
5	Докер. Подпомага работата на челния товарач в трюма - 6 ч	15,27	11,46	0,19	6,00
6	Водач на челен товарач в трюма. Избутва фосфорита от периферията към центъра на хамбара - 6 часа	13,73	10,30	0,16	6,00
	μ	15,777	10,195	0,175	6,000



изчисл. концентр.					
X			Y		
n1	6		n2	6	
$\mu_1=x$	0,000		$\mu_2=Y$	0,000	
\bar{X}	2,133	$(x_i - \bar{x})^2$	\bar{Y}	15,777	$(y_i - \bar{y})^2$
	2,670	0,28801		17,370	2,5387
	2,510	0,14188		15,580	0,0387
	2,410	0,07654		14,790	0,9735
	2,860	0,52804		17,920	4,5939
	1,280	0,72818		15,270	0,2567
	1,070	1,13068		13,730	4,1888
	$\Sigma(x_i - \bar{x})^2$	2,89333		$\Sigma(y_i - \bar{y})^2$	12,590
	n1-1	5		n2-1	5
	s_1^2	0,57867		s_2^2	2,51807
	s_1	0,76070		s_2	1,58684
	T	$\bar{X} - \bar{Y}$	\sqrt{A}	\sqrt{B}	
	-18,991	-13,643	5,477	3,935	
	$\alpha = 1 - P$	0,1			
а) Двустранен критерий	$\alpha/2$	0,05	P	0,9	
Ho:	$\mu_1 = \mu_2$	$ T \leq t$	$v = n_1 + n_2 - 2$	10	
H1:	$\mu_1 \neq \mu_2$	$ T > t$	от табл. I-1	$v=10; \alpha/2=0,05$	
	t		T	Ho се отхвърля: $\mu_1 = \mu_2$	
	2,228	<	-18,99	Приема се H1: $\mu_1 \neq \mu_2$	
б) Едностраниен критерий	α	0,1	P	0,9	
Ho:	$\mu_1 \leq \mu_2$	$T < t$	$v = n_1 + n_2 - 2$	10	
H1:	$\mu_1 > \mu_2$	$T > t$	от табл. I-1	$v=10; \alpha=0,1$	
	t		T	Ho не се отхвърля: $\mu_1 \leq \mu_2$	
	1,812	>	-18,99		

Фиг.5. Проверка на хипотези. Класове X-Y

3) Резултати

- нормирано разстояние между математическите очаквания

$$T = \frac{(\bar{x} - \bar{y})}{\sqrt{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}} \sqrt{\frac{n_1 \cdot n_2 (n_1 + n_2 - 2)}{n_1 + n_2}} = -18,991$$

Квантили на разпределенията на Стюдънт (Табл.4 [4])

При двустранен критерий $t\left(v, \frac{\alpha}{2}\right) = t(10; 0,05) = 2,228$

При едностраниен критерий $t(v, \alpha) = t(10; 0,1) = 1,812$

4) Проверка на хипотезите

- при двустранен критерий $|T| = 18,991 > t(10; 0,05) = 2,228$. Нулевата хипотеза се отхвърля. Приема се алтернативната $H_1: \mu_x \neq \mu_y$

- при едностраниен критерий $T = -18,991 < t(10; 0,1) = 1,812$. Нулевата хипотеза не се отхвърля – приема се $H_0: \mu_1 < \mu_2$

Заклучение: Имаме основание да приемем, че математическите очаквания μ_x и μ_y се различават съществено, а математическото очакване на концентрацията в класа, съответстващ на Наредба 13 (X) е значително по-малък от математическото очакване за втория клас (Y). Следователно може да се приеме предположението за наличие на два класа:

$$N_x(\mu_x, s_x) = N_x(2,133, 0,76); \quad N_y(\mu_y, s_y) = N_y(15,777, 1,587)$$

6.2.2. Проверка на хипотези за различия на математическите очаквания на групата от данни, съответстващи на Наредба 13 (x) и междинните данни от предполагаемия клас и „допустим, но опасен“. (z)

1. Числени данни на оценките на параметрите на разпределенията

$$\text{Клас X: } \mu_x = \mu_1 = \bar{x} = 2,133; \quad s_x^2 = 0,579; \quad s_x = 0,76; \quad n_x = 6;$$

$$\text{Клас Z: } \mu_z = \mu_3 = \bar{z} = 5,603; \quad s_z^2 = 9,852; \quad s_z = 3,138; \quad n_z = 4;$$

2. Хипотези:

а) двустранен критерий: $H_0: \mu_x = \mu_z$; $H_1: \mu_x \neq \mu_z$

б) едностранен критерий: $H_0: \mu_x \leq \mu_z$; $H_1: \mu_x > \mu_z$

Подробната реализация на алгоритмите за проверка на хипотезите за едностранен и двустранен критерий при предположение, че дисперсиите са неизвестни, но равни и представени на фиг.6.

3. Резултати:

- нормирано разстояние между математическите очаквания $T = -2,668$

- квантили на разпределението на Стюдънт (табл.4 [4]),

при двустранен критерий $t\left(v, \frac{\alpha}{2}\right) = t(8; 0,05) = 2,306$

при едностранен критерий $t(v, \alpha) = t(8; 0,1) = 1,86$

4. Проверка на хипотезите

- при двустранен критерий $|T| = 2,668 > t(8; 0,05) = 2,306$

Нулевата хипотеза се отхвърля. Приема се алтернативната $H_1: \mu_x \neq \mu_z$

- при едностранен критерий $T = -2,668 \leq t(8; 0,05) = 1,86$

Нулевата хипотеза не се отхвърля – приема се $H_0: \mu_x \leq \mu_z$

Заклучение: Имаме основание да приемем, че математическите очаквания μ_x и μ_z се различават съществено, а математическото очакване на концентрацията в класа, съответстващ на Наредба 13 (X) е по-малко от математическото очакване на данните от междинния клас (Z). Следователно може да се приеме предположението за различие на двата класа:

$$N_x(\mu_x, s_x) = N_x(2,133, 0,76); \quad N_z(\mu_z, s_z) = N_z(5,603, 3,138)$$

(X)

табл.1 - I-ви клас - Съответствие на Наредба №13/30.01.2004						
№	Работно място и изпълнител	изчисл. концентр.	средно прет. по	неопределеност	пределна стойност	
1	Сигналчик. От палуата дава указания на краниста. Продължителност 6 часа	2,67	2,01	0,04	3,50	
2	Кранист.Проверка, настройка и управление на конвейра (в трюма) Продължителност 4 часа	2,51	1,26	0,02	3,50	
3	Кранист. Разтоварване по указание на сигналчика. Продължителност 4 часа.	2,41	1,21	0,02	3,50	
4	Обходчик-до конвейра. Следи работата на трансп.лента - 5 часа	2,86	1,79	0,03	4,00	
5	Обходчик при пресипване. Следи работата на лентата - 5 часа	1,28	1,04	0,02	3,50	
6	Всички работници изчакват в битовата зала - 4 часа	1,07	0,53	0,01	3,50	
		μ	2,133	1,307	0,023	3,583



(Z)

табл.3 - III-ти клас - "допустим, но опасен"						
№	Работно място и изпълнител	изчисл. концентр.	средна по	неопределеност	пределна стойност	
1	Обходчик-Осигурява функционирането на трансп.лента (при пресипване) - 6,5 часа	9,71	7,89	0,14	6,00	
2	Кранист, сигналчик, обходчик, водач на товарач изчакват в битовата зала - 4 часа	6,40	3,20	0,05	6,00	
3	Докер. Подпомага работата на челния товарач в трюма на кораба - 6 часа	3,25	2,44	0,04	3,50	
4	Водач на челен товарач в трюма. По указания на сигналчика избутва фосфорита от периферията към центъра на трюма - 6 часа	3,05	2,29	0,04	3,50	
		μ	5,603	3,955	0,068	4,750



изчисл. концентр.					
X			Z		
n1	6		n2	4	
μ1=x	0,000		μ2=Z	0,000	
\bar{x}	2,133	(xi - x)^2	\bar{z}	5,603	(zi - z)^2
	2,670	0,28801		9,710	16,8716
	2,510	0,14188		6,400	0,6360
	2,410	0,07654		3,250	5,5343
	2,860	0,52804		3,050	6,5153
	1,280	0,72818			
	1,070	1,13068			
	$\sum(xi - x)^2$	2,89333		$\sum(zi - z)^2$	29,557
	n1-1	5		n2-1	3
	s1^2	0,57867		s2^2	9,85236
	s1	0,76070		s2	3,13885
	T	X - Z	√A	√B	
	-2,668	-3,469	4,382	5,697	
	α = 1 - P	0,1			
а) Двустранен критерий		α/2	0,05	P	0,9
Ho:	μ1 = μ2	T ≤ t	v = n1+n2 - 2	8	
H1:	μ1 ≠ μ2	T > t	от табл. I-1	v=8; α/2=0,05	
	t		T	Но се отхвърля: μ1 = μ2	
	2,306	<	-2,67	Приема се H1: μ1 ≠ μ2	
б) Едностраниен критерий		α	0,1	P	0,9
Ho:	μ1 ≤ μ2	T < t	v = n1+n2 - 2	8	
H1:	μ1 > μ2	T > t	от табл. I-1	v=8; α=0,1	
	t		T	Но не се отхвърля: μ1 ≤ μ2	
	1,86	>	-2,67		

Фиг.6. Проверка на хипотези. Класове X-Z

6.2.3. Проверка на хипотези за различие на математическите очаквания на групите от данни, несъответстващи на Наредба 13 (Y) и междинните данни от предполагаемия клас „допустим, но опасен“ (Z).

1) Числени стойности на оценките на параметрите на разпределенията.

Клас Y: $\mu_y = \mu_1 = \bar{y} = 15,777$; $s_y^2 = 2,518$; $s_y = 1,587$; $n_y = 6$

Клас Z: $\mu_z = \mu_2 = \bar{z} = 5,603$; $s_z^2 = 9,852$; $s_z = 3,138$; $n_z = 4$

2) Хипотези:

а) двустранен критерий: $H_0: \mu_y = \mu_z$; $H_1: \mu_y \neq \mu_z$

б) едностранен критерий: $H_0: \mu_y \leq \mu_z$; $H_1: \mu_y > \mu_z$

Подробната реализация на алгоритмите за проверка на хипотезите при едностранен и двустранен критерий при предположение, че дисперсиите са неизвестни, но равни е представена на фиг.7.

3) Резултати

- нормално разстояние между математическите очаквания $T = 6,867$

- квантили на разпределението на Стюдънт (Табл.4 [11])

При двустранен критерий $t\left(v, \frac{\alpha}{2}\right) = t(8; 0,05) = 2,306$

При едностранен критерий $t(v, \alpha) = t(8; 0,1) = 1,86$

4) Проверка на хипотезите

- при двустранен критерий $|T| = 6,867 > t(8; 0,05) = 2,306$. Нулевата хипотеза се отхвърля- приема се $H_1: \mu_y \neq \mu_z$

- при едностранен критерий $T = 6,867 > t(8; 0,1) = 1,86$. Нулевата хипотеза се отхвърля – приема се $H_1: \mu_y > \mu_z$.

Заключение: Имаме основание да приемем, че математическите очаквания μ_y и μ_z се различават съществено, а математическото очакване на концентрацията в класа, несъответстващ на Наредба 13 (Y) е по-голямо от математическото очакване на данните от междинния клас (Z).

Следователно може да се приеме предположението за различие на разпределенията за двата класа $N_y(\mu_y, s_y) = N_y(15,777, 1,57)$; $N_z(\mu_z, s_z) = N_z(5,603, 3,138)$.

(Y)

табл.2 - II-ри клас - Не съответствие на Наредба №13/30.01.2004					
№	Работно място и изпълнител	изчисл. концентр.	средно прет. по	неопределенност	пределна стойност
1	Сигналчик. От палуата дава указания на краниста при разтоварване 6 часа	17,37	13,02	0,24	6,00
2	Кранист.Проверка, настройка и управление на конвейра (в трюма) - 4 часа	15,58	7,79	0,14	6,00
3	Кранист. Разтоварва насипни товари от трюма по указание на сигналчика - 4 часа.	14,79	7,40	0,13	6,00
4	Обходчик-Осигурява функционирането на трансп.лента (до конвейра) - 5 часа	17,92	11,20	0,19	6,00
5	Докер. Подпомага работата на челния товарач в трюма - 6 часа	15,27	11,46	0,19	6,00
6	Водач на челен товарач в трюма. Избутва фосфорита от периферията към центъра на хамбара - 6 часа	13,73	10,30	0,16	6,00
	μ	15,777	10,195	0,175	6,000



(Z)

табл.3 - III-ти клас - "допустим, но опасен"					
№	Работно място и изпълнител	изчисл. концентр.	средна по	неопределеност	пределна стойност
1	Обходчик-Осигурява функционирането на трансп.лента (при пресипване) - 6,5 часа	9,71	7,89	0,14	6,00
2	Кранист, сигналчик, обходчик, водач на товарач изкаква в битовата зала - 4 часа	6,40	3,20	0,05	6,00
3	Докер. Подпомага работата на челния товарач в трюма на кораба - 6 часа	3,25	2,44	0,04	3,50
4	Водач на челен товарач в трюма. По указания на сигналчика избутва фосфорита от периферията към центъра на трюма - 6 часа	3,05	2,29	0,04	3,50
	μ	5,603	3,955	0,068	4,750

		изчисл. концентр.			
		Y		Z	
n1	6	n2	4		
μ1=y	0,000	μ2=Z	0,000		
\bar{y}	15,777	\bar{z}	5,603	$(y_i - y)^2$	$(z_i - z)^2$
	17,370		9,710	2,53871	16,8716
	15,580		6,400	0,03868	0,6360
	14,790		3,250	0,97351	5,5343
	17,920		3,050	4,59388	6,5153
	15,270			0,25671	
	13,730			4,18884	
	$\sum(y_i - y)^2$		$\sum(z_i - z)^2$	12,59033	29,557
	n1-1		n2-1	5	3
	s1^2		s2^2	2,51807	9,85236
	s1		s2	1,58684	3,13885
	T	Y - Z	√A	√B	
	6,867	10,174	4,382	6,492	
	$\alpha = 1 - P$	0,1			
а) Двустранен критерий	$\alpha/2$	0,05	P	0,9	
Ho:	$\mu_1 = \mu_2$	T ≤ t	v = n1+n2 - 2	8	
H1:	$\mu_1 \neq \mu_2$	T > t	от табл.1-1	v=8; α/2=0,05	
	t		T	Но се отхвърля: $\mu_1 = \mu_2$	
	2,306	<	6,87	Приема се H1: $\mu_1 \neq \mu_2$	
б) Едностраниен критерий	α	0,1	P	0,9	
Ho:	$\mu_1 \leq \mu_2$	T < t	v = n1+n2 - 2	8	
H1:	$\mu_1 > \mu_2$	T > t	от табл.1-1	v=8; α=0,1	
	t		T	Но се отхвърля: $\mu_1 \leq \mu_2$	
	1,86	<	6,87	Приема се H1: $\mu_1 > \mu_2$	

Фиг.7. Проверка на хипотези. Класове Z-Y

6.3. Реализация на алгоритмите за проверка на хипотези за равенство на две дисперсии

Тази задача от изследването, осъществена по алгоритъма за проверка на хипотези (5.3) има за цел създаването на условия за определяне на статистически обосновани граници между класовете на състояние при повече от два класа. Проверката на хипотезите за равенство на дисперсиите при три класа на състояние, каквато е решаваната от нас задача ще се осъществява за всяка двойка от данни $(X Y)$, $(X Z)$, $(Y Z)$.

6.3.1. Проверка на хипотези за равенство на дисперсиите на групите от данни, съответстващи (X) и несъответстващи (Y) на Наредба 13/ДВ. бр. 8/30.01.2004 г.

1) Числени стойности на оценките на параметрите на разпределенията

Клас X : $\mu_x = \bar{x} = 2,133$; $s_x^2 = 0,579$; $s_x = 0,76$; $n_x = 6,0$

Клас Y : $\mu_y = \bar{y} = 15,777$; $s_y^2 = 2,518$; $s_y = 1,587$; $n_y = 6,0$

За да бъдат изпълнени условията за прилагане на критерия на Фишер (5.3.4), означаваме като s_1^2 по-голямата дисперсия (в случая това е s_y^2).

2) Хипотези

Ако при формулирането на хипотезите се използва горното условие на критерия на Фишер, трите класически случая на проверка на хипотези ще бъдат сведени до проверка на една хипотеза [4]:

$$H_0 : \sigma_1^2 = \sigma_2^2 ; H_1 : \sigma_1^2 > \sigma_2^2$$

3) Съставя се дисперсното отношение

$$F = \frac{s_1^2}{s_2^2} = \frac{2,518}{0,579} = 4,349$$

като в числителя се поставя по-голямата дисперсия $s_1^2 > s_2^2$.

Изчислява се броят на степените на свобода:

- за числителя $v_1 = n - 1 = 5$

- за знаменателя $v_2 = 5$

От таблицата за разпределение на Фишер (Табл. 5, Приложение [4]) при зададено ниво на значимост $\alpha = 0,05$ се отчита критичната стойност за разпределението на Фишер-Снедекор

$$F(\alpha, v_1, v_2) = F(0,05 ; 5; 5) = 5,05$$

4) Проверка на хипотезите

$$F = 4,349 < F(0,05 ; 5; 5) = 5,05$$

Нулевата хипотеза $H_0 : \sigma_1^2 = \sigma_2^2$ не се отхвърля (приема се $\sigma_1^2 = \sigma_2^2$), което означава, че $\sigma_x^2 = \sigma_y^2$.

6.3.2. Проверка на хипотези за равенство на дисперсиите на групата от данни, съответстващи на Наредба 13 (X) и на междинните данни от предполагаемия клас „допустим, но опасен“ (Z) .

1) Числени стойности на оценките на параметрите на разпределенията

Клас X: $\mu_x = \bar{x} = 2,133$; $s_x^2 = 0,579$; $s_x = 0,76$; $n_x = 6$

Клас Z: $\mu_z = \bar{z} = 5,603$; $s_z^2 = 9,852$; $s_z = 3,139$; $n_z = 4$

Означаваме $s_1^2 = s_z^2$; $s_2^2 = s_x^2$

2) Хипотези

$H_0: \sigma_1^2 = \sigma_2^2$; $H_1: \sigma_1^2 > \sigma_2^2$

3) Дисперсно отношение

$$F = \frac{s_1^2}{s_2^2} = \frac{9,852}{0,579} = 17,01$$

Степени на свобода:

- за числителя $v_1 = n_1 - 1 = 3$

- за знаменателя $v_2 = n_2 - 1 = 5$

От таблицата за разпределение на Фишер (Табл. 5, Приложение [4]) при зададено ниво на значимост $\alpha = 0,05$ и степени на свобода $v_1 = 3$, $v_2 = 5$ се отчита критичната стойност

$$F(\alpha, v_1, v_2) = F(0,05; 3; 5) = 5,41$$

4) Проверка на хипотезите

$$F = 17,1 > F(0,05; 3; 5) = 5,41$$

Нулевата хипотеза $H_0: \sigma_1^2 = \sigma_2^2$ се отхвърля. Приема се хипотезата $H_1: \sigma_1^2 > \sigma_2^2$, следователно $\sigma_z^2 > \sigma_x^2$

6.3.3. Проверка на хипотези за равенство на дисперсиите на групата от данни, несъответстващи на Наредба 13 (Y) и на междинните данни от предполагаемия клас „допустим, но опасен“ (Z).

1) Числени стойности на оценките на параметрите на разпределенията

Клас Y: $\mu_y = \bar{y} = 15,77$; $s_y^2 = 2,518$; $s_y = 1,587$; $n_y = 6$

Клас Z: $\mu_z = \bar{z} = 5,603$; $s_z^2 = 9,852$; $s_z = 3,139$; $n_z = 4$

Означаваме $s_1^2 = s_z^2$; $s_2^2 = s_y^2$

2) Хипотези

$H_0: \sigma_1^2 = \sigma_2^2$; $H_1: \sigma_1^2 > \sigma_2^2$

3. Дисперсно отношение

$$F = \frac{s_1^2}{s_2^2} = \frac{9,852}{2,518} = 3,91$$

Степени на свобода:

- за числителя $v_1 = 3$

- за знаменателя $v_2 = 5$

От таблицата за разпределение на Фишер (Табл. 5, Приложение [4]) при зададено ниво на значимост $\alpha = 0,05$ и степени на свобода $\nu_1 = 3$, $\nu_2 = 5$ се отчита критичната стойност

$$F(\alpha, \nu_1, \nu_2) = F(0,05; 3; 5) = 5,41$$

4) Проверка на хипотезите

$$F = 3,91 < F(0,05; 3; 5) = 5,41$$

Нулевата хипотеза $H_0: \sigma_x^2 = \sigma_y^2$ не се отхвърля. Приема се хипотезата за равенство на двете дисперсии $\sigma_x^2 \approx \sigma_y^2$.

Заклучение:

Резултатите от проведените проверки на хипотези за стойностите и разположенията на параметрите на разпределенията на изследваните три групи от експериментални данни ни дават основание да не отхвърляме твърдението за съществуването на три едномодални закона на разпределение:

$$N_x(T_x, \sigma_x) = N_x(2,133; 0,76);$$

$$N_y(T_y, \sigma_y) = N_y(15,777; 1,587);$$

$$N_z(T_z, \sigma_z) = N_z(5,603; 3,139). \quad (25)$$

Като илюстрация на резултатите от проверката на хипотези са представени разпределенията на данните от измерванията с концентрация на прах от фосфорит във въздуха съответстващи (фиг. А) и несъответстващи (фиг. Б) на изискванията на Наредба 13/ДВ. Бр.8/ 30.01.2004г. Специално внимание е обърнато на разположенията на тези разпределения по отношение на априорно зададените гранични стойности $\mu_0 = 3,5$ и $\mu_0 = 6,5$

От разпределението на данните, отговарящи на изискванията на Наредба 13, фиг.А се вижда, че с увеличаване на граничните стойности μ_0 , мощността на критерия $H_0: \mu_x \leq \mu_0$ расте (при $\mu_0 = 3,5$, $T = -4,4$, $P = 0,9$; при $\mu_0 = 6$, $T = -12$, $P = 0,99$) като след $\mu_0 > 5$ може да се приеме нулевата хипотеза с максимална вероятност ($P \approx 1$; $\alpha \approx 0$).

От разпределението на данните, неотговарящи на изискванията на Наредба 13, фиг.Б се вижда, че за тези данни мощността на критерия $H_0: \mu_y \geq \mu_0$ е максимално възможна с $P \approx 1$ ($\alpha \approx 0$) до граница $\mu_0 = 10$. Това отговаря на една зона на изменение на граничните стойности на контролирания показател (концентрация на фосфорит) $5 < \mu_0 < 10$, в която не може да се вземе никакво управляващо решение. Едва след гранична стойност $\mu_0 > 10$ започва едно намаление на мощността на критерия $H_0: \mu_y \geq \mu_0$, което прави възможно и необходимо решаването на задачата за проверка на хипотези.

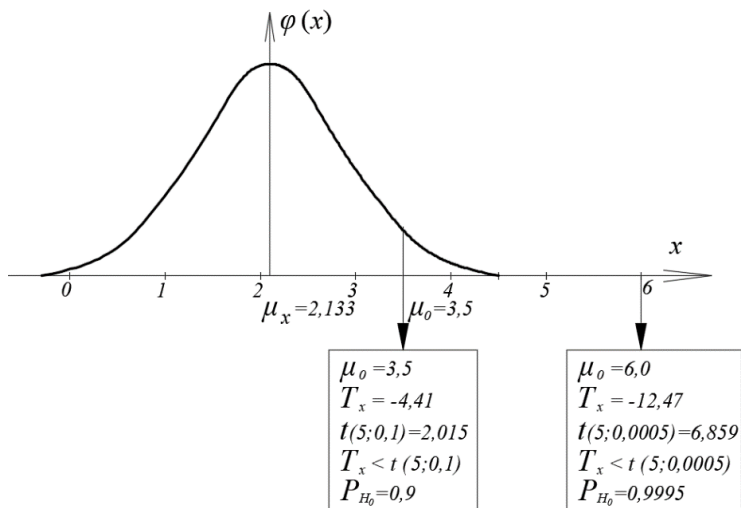
Наличието на широкия интервал на изменение на зоната на нечувствителност на действията за откриване и ликвидиране на екологично опасната ситуация поставя въпроса за целите на работа на участниците в процеса на разтоварването. Това поставя задачата или за преразпределение на граничните стойности или въвеждането на един междинен клас, през който работниците да бъдат по-добре обезопасени (фиг. В).

Тъй като броят на експериментите, проведени в експлоатационни условия не е достатъчен за приемане на генерални и еднозначни доказателства, обръщаме внимание на устойчивостта на процедурите за проверките към нарушаване на предпоставките за нормалност на разпределенията. Вече описаните и приложени процедури за проверка на хипотезите за различие на математическите очаквания, използващи разпределението на Стюдънт са относително нечувствителни към нарушаване на предположенията за нормалност [10]. С увеличаване на обема на данните и с използването на еднакви обеми на сравняваните разпределения процедурите стават все по-устойчиви. Причината за това е, че разпределението на Стюдънт асимптотично се стреми към нормалното, когато $n \rightarrow \infty$. Същото се отнася и за процедурите за проверка на хипотезите за равенство на математическите очаквания, основаващи се на разпределението на Фишер [9, 34]. Проверките на хипотезите за равенство на дисперсиите на трите групи от данни (X, Z, Y) показаха, че нулевите хипотези за равенство на дисперсиите на класовете (X→Y) и (Z→Y) не се отхвърлят ($\sigma_x^2 \approx \sigma_y^2$ и $\sigma_z^2 \approx \sigma_y^2$), а за класовете (X→Z) се приема алтернативната хипотеза $H_1: \sigma_z^2 > \sigma_x^2$ (процедура 6.3.2 и уравнения 25). Препоръките към експериментаторите при съмнения относно нормалността на разпределенията са да се използват голям брой на наблюденията, чиито обем за отделните извадки да бъде един и същ [4,5,10,11,12,14,15,16]. С това ще се изпълнят условията за асимптотично приближаване на емпиричните разпределения към нормалното, с което ще се повишат възможностите за използване на резултатите от проверките на хипотези в задачите и стандартите за контрол на качеството.

$$H_0: \mu_x \leq \mu_0, \quad n_x = 6; \quad \bar{x} = 2,133, \quad s_x = 0,76$$

$$T_x = \frac{(\bar{x} - \mu_0)\sqrt{n}}{s_x}$$

H_0 не се отхвърля при $T < t(v, \alpha)$

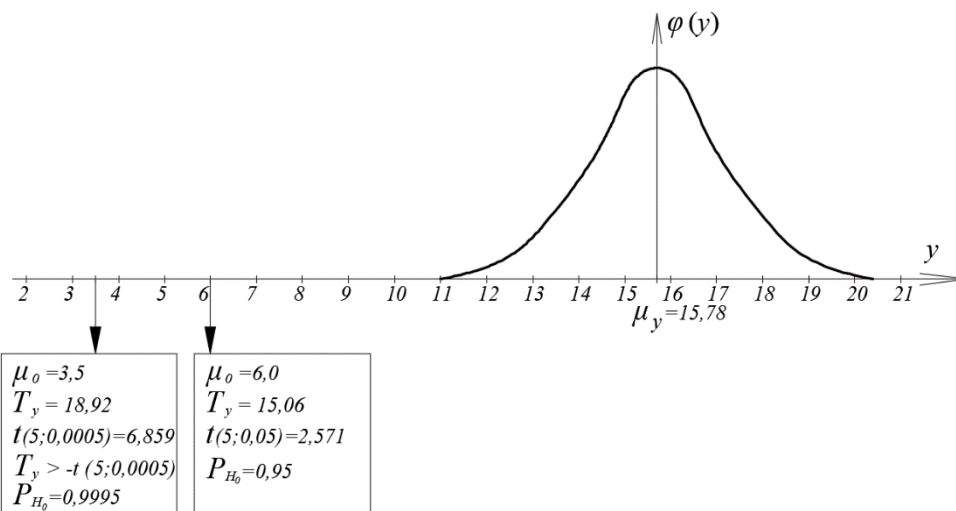


Фиг.А. Разпределение и проверка на хипотези за данните X - отговарящи на Наредба 13.

$H_0: \mu_y \geq \mu_0, n_y = 6; \bar{y} = 15,78, s_y = 1,59$

$$T_y = \frac{(\bar{y} - \mu_0)\sqrt{n}}{s_y}$$

H_0 не се отхвърля при $T > -t(v, \alpha)$

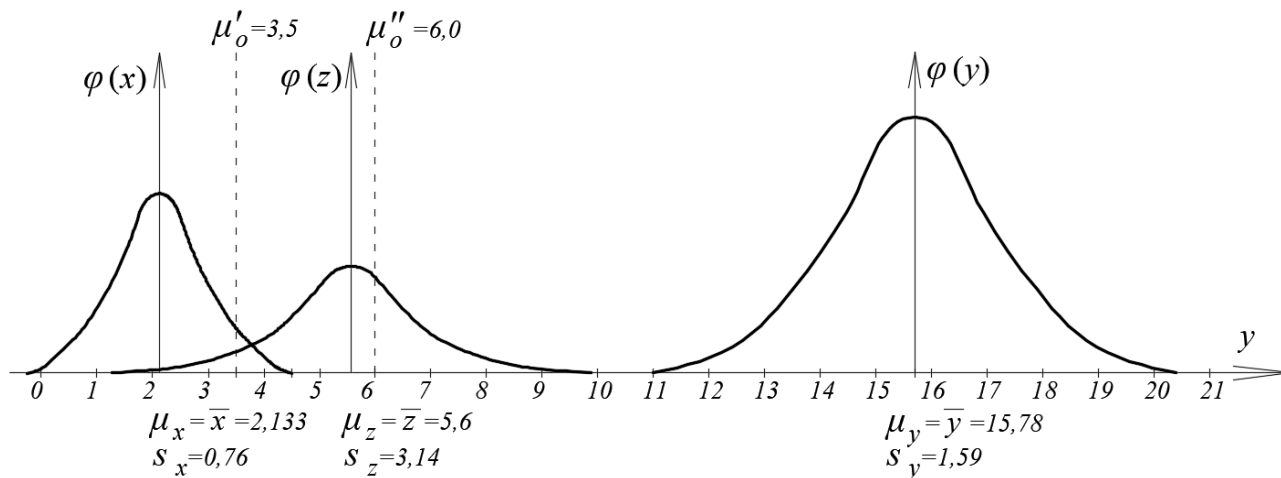


Фиг.Б. Разпределение и проверка на хипотеза за данните Y - неотговарящи на Наредба 13.

X - данни, отговарящи на Наредба 13

Y - данни, неотговарящи на Наредба 13

Z - предполагаем клас „допустим, но опасен“



Фиг.В. Разпределение на предполагаемите три класа (X, Y, Z) и разположение на техните центрове по отношение на граничните стойности

VII. Определяне на статистически обосновани гранични стойности между класовете на състояние при решаване на задачата за диагностичен контрол по косвени признаци

Като имаме предвид вече установената последователност на разположението на центровете на разпределенията $\mu_x < \mu_z < \mu_y$ си поставяме в следващата част от изследването задачата за определяне на граничните стойности на контролирания едномерен показател (например прах от фосфорит) между последователните класове на състояние. В теоретичните източници или в стандартите за контрол на качеството няма единно мнение по този въпрос. В някои източници (например в Наредба 13/ДВ, бр. 8/30.01.2004 г. и следващите изменения към нея до 2021 г.) се говори за две гранични стойности между два съседни класа. В стандартите и методиките за контрол не се среща задачата за определяне на граничните стойности между влошаващите се класове на състоянието, когато техния брой е по-голям от два (например три). Поради това ще потърсим теоретично обосноваване и първоначално практическо решение на тези задачи, като използваме вероятностния подход за оптимално стохастическо управление, намерил приложение в техническата диагностика [10,11,16,30].

7.1. Изходни данни.

Теоретичното решение на задачата за определяне на най-добрата (статистически оптимална) стойност на разделящата граница между два пресичащи се класа е предложено в трудовете на А. Недев (например [1]). По принцип задачата има решение за повече от два класа, но тук се приема подходът за определяне на границата между всяка двойка от класове V_1 и V_2 , условно наречени „добър и недобър“. Задачата е: като се използват измерените стойности на контролирания параметър (например концентрацията на прах) да се определи такава стойност на този показател t_0 , разделяща двата класа с най-малка грешка (заштрихованата площ на фиг. 8).

По вече разгледания подход сме определили параметрите на разпределенията на признака за всеки от класовете на състояние.

Клас X :

Плътност на разпределение на показателя (t) в областта X , която при малък обем на данните се описва със закона на Стюдънт, а с увеличаване на обема асимптотично клони към нормален закон.

$$f(t/x) = \frac{1}{\sigma_x \sqrt{2\pi}} \exp \left[-\frac{(t-\mu_x)^2}{2\sigma_x^2} \right] \quad (26)$$

$$\mu_x = 2,133; \quad \sigma_x^2 = 0,579; \quad \sigma_x = 0,76$$

Клас Y :

$$f(t/y) = \frac{1}{\sigma_y \sqrt{2\pi}} \exp \left[-\frac{(t-\mu_y)^2}{2\sigma_y^2} \right] \quad (27)$$

$$\mu_y = 15,777; \quad s_y^2 = 2,518; \quad s_y = 1,587$$

Клас Z:

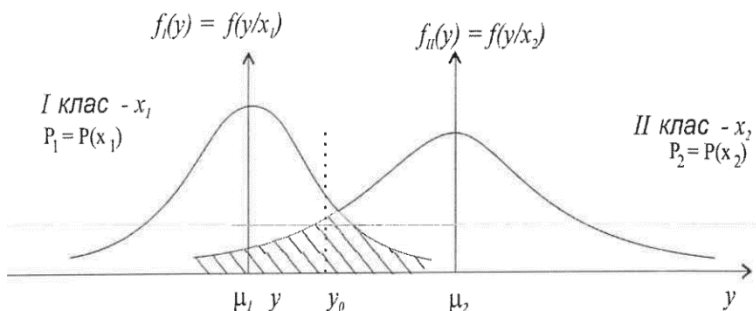
$$f(t/z) = \frac{1}{\sigma_z \sqrt{2\pi}} \exp \left[-\frac{(t-\mu_z)^2}{2\sigma_z^2} \right] \quad (28)$$

$$\mu_z = 5,603; \quad s_z^2 = 9,852; \quad s_z = 3,139$$

7.2. Теоретичен модел за определяне на граничните стойности между последователните класове.

При разработване на модела използваме подробно изложените във втора глава теоретични основи на методите за разпознаване на образи.

Като имаме предвид задачата, която предстои да бъде решавана, приемаме, че състоянието, представляващо пълна група от две несъвместими събития (например добро и недобро) се описва с помощта на един контролируем показател (например концентрацията на прах). При това законите на разпределение на контролируемия (диагностичен) признак, получени в резултат на предварителни експерименти се пресичат (фиг. 8).



Фиг. 8. Разпределение на диагностичния признак при два класа на състояние

Тази двуалтернативност означава, че при три последователни класа X, Z, Y ще се търсят две гранични стойности на параметъра t (t_{0xz}, t_{0zy}), т.е. в общия алгоритъм ще търсим оптималната разделяща граница между две групи от данни, означени като V₁ и V₂.

В общия случай предполагаме, че в резултат на предварително проведени надеждности или общоекологични изследвания на обекта (технологическия процес) са получени оценките на априорните вероятности за поява на всеки от класовете P_I и P_{II}.

Задачата, която си поставяме е да намерим най-добрата в предварително избран смисъл оценка на разделящата граница между два класа, която в случая ще бъде търсената стойност на t₀. Отчитайки възможността за достатъчно широк кръг от алтернативи, ще ограничим понятието „оптималност“ в Байесов смисъл [10,11,12]. В този случай класифициращото решение се свежда до приемането като по-достоверна на онази хипотеза, чиято апостериорна вероятност е по-голяма, т.е.:

$$t \in V_1, \text{ ако } P_I \cdot f_I(t) > P_{II} \cdot f_{II}(t); \quad (i, j = 1, 2, \dots, i, j) \quad (29)$$

Или приведено към една от двете алтернативи :

$$t \in V_1, \text{ ако } P_I \cdot f_I(t) > P_{II} \cdot f_{II}(t) \quad (30)$$

$$t \in V_2, \text{ ако } P_I \cdot f_I(t) < P_{II} \cdot f_{II}(t)$$

Търсената стойност на разделящата граница t_0 е решението на уравнението:

$$P_I \cdot f_I(t) = P_{II} \cdot f_{II}(t) \quad (31)$$

намиращо се между най-вероятните стойности μ_1 и μ_2 на конкретните сравнявани едномодални разпределения (уравнения 25, 26, 27 и 28). Решението се търси при приемане на предположението за нормални или асимптотично доближаващи се до нормални на всяко от разпределенията $f(t/V_1)$ и $f(t/V_2)$:

$$f(t/V_1) = \frac{1}{\sigma_1 \sqrt{2\pi}} \exp \left[-\frac{(t-\mu_1)^2}{2(\sigma_1)^2} \right]; \quad f(t/V_2) = \frac{1}{\sigma_2 \sqrt{2\pi}} \exp \left[-\frac{(t-\mu_2)^2}{2(\sigma_2)^2} \right] \quad (32)$$

След заместване на (22) в (21) получаваме:

$$\frac{P_I}{\sigma_1} \exp \left[\frac{(t-\mu_1)^2}{2 \cdot \sigma_1^2} \right] = \frac{P_{II}}{\sigma_2} \exp \left[\frac{(t-\mu_2)^2}{2 \cdot \sigma_2^2} \right], \quad (33)$$

което след логаритмуване приема вида:

$$\left[\frac{(t-\mu_1)^2}{2 \cdot \sigma_1^2} \right] - \left[\frac{(t-\mu_2)^2}{2 \cdot \sigma_2^2} \right] = \ln \frac{P_{II} \sigma_1}{P_I \sigma_2}, \quad (34)$$

от където следва:

$$\frac{t^2}{2} \left(\frac{1}{\sigma_2^2} - \frac{1}{\sigma_1^2} \right) + t \left(\frac{\mu_1}{\sigma_1^2} - \frac{\mu_2}{\sigma_2^2} \right) + \left(\frac{\mu_2^2}{2\sigma_2^2} - \frac{\mu_1^2}{2\sigma_1^2} \right) = \ln \frac{P_{II} \sigma_1}{P_I \sigma_2} \quad (35)$$

Търсената стойност $t = t_0$ е този корен на квадратното уравнение (35), който е разположен между центровете на двете разпределения μ_1 и μ_2 .

Може да бъде потърсено и приблизително аналитично решение на горното уравнение, ако се приеме, че $\sigma_1^2 = \sigma_2^2 = \sigma^2$. Този подход се използва в задачите за разпознаване и класификация на състоянията по многомерния линеен дискриминант на Фишер, който се получава при равенство на ковариационните матрици, чиито едномерен аналог са дисперсиите [4,5,10]. В този случай уравнение (35) става линейно, а решението за оптималната разделяща граница е:

$$t_0 = \frac{\sigma^2}{\mu_2 - \mu_1} \ln \frac{P_I}{P_{II}} + \frac{\mu_1 + \mu_2}{2} \quad (36)$$

Това означава, че наблюдението t следва да бъде отнесено към клас V_1 , ако то е по-малко от граничната стойност t_0 , получена от (36).

Стеснявайки още повече условието на задачата, можем да предположим, че априорните вероятности на двата класа са равни (или неизвестни) $P_I = P_{II} = P$.

В този случай правилото за приемане на решение за разделящата граница приема вида на известната минимаксна стратегия [10,12,15] :

$$t_0 = \frac{\mu_1 + \mu_2}{2} \quad (37)$$

На базата на последователното прилагане на процедурите (37), (36), (35) се предлага една адаптивна стратегия за достигането на най-добрата стойност на разделящата граница, при която се започва от минимаксната оценка (37). Очевидно е, че отчитането на априорните вероятности и реалните характеристики на статистиката води до уточняване на разделящата граница чрез създаване на възможности за използване на процедурите (36) и (35).

Полезността на предложения адаптивен подход за търсене на оптимално в статистически смисъл решение е очевидна и тя следва от факта, че и в най-добрия случай едва ли биха се намерили диагностични показатели и способности за тяхното измерване, които да позволяват абсолютното разделяне на съседните класове на състояние.

7.3. Резултати.

Следвайки приетата адаптивна стратегия за уточняване на стойностите на границите между класовете на състояние в зависимост от пълнотата на данните за разпределенията, представяме резултатите в последователността на алгоритмите (37), (36), (35). Конкретните уравнения и решения се получават след заместване на изходните данни (уравнения 26,27,28) в алгоритмите (35), (36), (37).

а) граници между класовете по правилото на минимакса:

$$t_{0x,z} = 3,9; \quad t_{0z,y} = 10,7$$

б) граници при линейна стратегия на Байес ($\sigma^2 = 4,136$)

$$t_{0x,z} = 3,0 \text{ (при } P_x = 0,4; P_z = 0,6); \quad t_{0z,y} = 11,5 \text{ (при } P_z = 0,6; P_y = 0,4)$$

в) уравнения на границите по квадратичен Байесов алгоритъм

$$\text{за определяне на } t_{0x,z}: \quad 0,815 t^2 + 3,115 t - 3,73 = 0$$

$$\text{за определяне на } t_{0z,y}: \quad 0,149 t^2 - 5,697 t + 47,11 = 0$$

Тук не представяме различните варианти на решенията, които са много чувствителни по отношение на параметрите на законите на разпределение. В резултат на вече установената чувствителност на методите за оценка и проверка на хипотези на параметрите на разсейването от общия обем на данните и разпределението им в различните класове са получени оценки на дисперсиите за междинния клас, които се отличават от другите дисперсии на порядък. За получаването на надеждни резултати, които да бъдат основа за бъдещи стандарти за управление на качеството, е необходимо провеждането на специализирани изпитания в широк мащаб от външни и управляващи условия. На базата на такава статистическа основа и с използването на посочените тук теоретични подходи може да се пристъпи към коригиране на стандартите за контрол на качеството.

Литература:

1. Каменова С. Замърсяване на въздуха и въздействие върху екосистемите. ТУ-Варна, 2001г. 273 с. ISBN 954-20-0156-8;
2. Консулова-Бакалова М. Автоматизация на инженерния труд. „Геа Принт“ Варна, 2021г., 237с. ISBN 978-619-184-047-2
3. Близнаков А. Модел за анализ на екологично опасните показатели при товаро-разтоварни дейности в морските пристанища. E-Jurnal "Морско право и Морска индустрия, бр2. 2024 г. ISSN: 2815-5130
4. Божанов Е, И, Вучков, Статистически решения в производството и научните изследвания. „Техника“, С.1979г., 409с.
5. Венцель Е.С. „Теория вероятностей“, „Наука“, М.1968;
6. Тъкър Х. Математическа статистика, Наука и изкуство, С.1968г.
7. Кендал М. Д, А.Стюарт. Статистически изводи и връзки. Наука. М.,1973 г.
8. Уилкс С. Математическа статистика. Наука.М.1976г.
9. Фишер Р.А. Статистически методи за изследователи. Гостиздат. М.1958г.
10. Атанасов А., Диагностика и управление на техническата експлоатация на КЕУ. Дисертация за научна степен д.т.н. ТУ-Варна 1989 г.
11. Недев А., Разпознаване на образи и оптимално стохастическо управление. Книга I-ва И.К.Геа-Принт. Варна. 2012 г., 345с.
12. Недев А., Разпознаване, диагностика, управление. Теоретични основи и приложни задачи. И.К.Геа-Принт. Варна. 2020 г., 352с.
13. Недев А., Диагностика и управление на техническата експлоатация на корабни енергетични уредби. И.К.Геа-Принт. Варна. 2019 г., 445с.
14. Антонов Г., Експертна система за оценка на техническото състояние на обекти от химическото оборудване. сп.Машиностроене и машинознание, бр.10, Кн.1, ISSN1312-8612, 2010 г.
15. Недев А., М. 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 VI.2009 г., 455 с.
16. Недев А., М. Бакалова, Г. Антонов, Б. Андреев, С.Сезгин, Д.Камберов. Разпознаване на образи и оптимално стохастическо управление. Кн.2. Приложение на методите за различаване на образи в управлението на стопански, биологични и обществени системи. Варна, И.К.Геа-Принт. Варна. 2012 г., 368с.
17. Недев А., А.Георгиева, Ал.Близнаков и др. Определяне на нивото на последствията от екологичните опасности на морските транспортни дейности. E-Journal. VFU, бр.10, 2023.
18. Чакър А., Р.Атанасов, Ал.Близнаков. Байесов подход за оценка на пасивните мерки за защита от пожар в помещение. E-Journal. VFU, бр.15, 2021.
19. Димитракиев Д., К.Донев, Я.Ганчева. Пристанища, И.К.Ларго Сити 001.Варна. 2014

20. Димитрова В., А.Недев. Сравнение на приложимостта на невронни мрежи при разпознаване на нефтопродукти. сп. Механика на машините. бр.77, 2008, ISSN0861-9727, стр.70-73
21. Консулова М., П.Наскова, Б.Малчева, Др.Пламенов. Комбиниране на статистически критерии при определяне на числеността на почвената микрофлора. сп. Ново знание, год.6, бр.4, 2017 г.
22. Консулова М., П.Наскова, Др.Пламенов., Б.Малчева. Математически модел за определяне на степента на влияние на различни физико-химични фактори върху числеността на общата микрофлора в антропогенни почви. сп. Ново знание, год.6, бр.4, 2017 г., стр.88-102.
23. Консулова М., П.Наскова, Др.Пламенов., Б.Малчева. Разпознаване и прогнозиране на почвена микробиологична активност по косвени признаци. сп.Почвознание, агрохимия и екология. бр.51, №3-4, стр.12-20
24. Консулова М., М.Бакалова. Възможности за разпознаване и представяне на силно размити данни от газов анализ. сп. Механика на машините. бр.57, 2005, ISSN0861-9727, стр.11-14
25. Консулова М., М.Бакалова, А.Недев, Параметрични и непараметрични методи за разпознаване на данни от устройство тип „електронен нос“. сп. Механика на машините. бр.77, 2008, ISSN0861-9727, стр.41-44
26. Наскова П. Моделиране на процесите в почвено-растителните екосистеми. И.К.Геа-Принт. Варна. 2020 г., 241с.
27. Наскова П. Моделиране на динамиката на органични вещества и тежки метали в почвата.И.К.Геа-Принт. Варна. 2024 г., 192с.
28. Наскова П. Математически модел за оценка на съдържанието на тежки метали в почвата по косвени признаци на растенията. сп. Ново знание, vol.6, №3, 2017 г., стр.149-160
29. Кендал М., А.Стьюарт. Теория распределений. Наука М.1966
30. Дуда Р., П.Харт. Распознавание образов и анализ сцен. Мир, Москва, 1976, с.510
31. Фуконага К. Введение в статическую теорию распознавания образов. Наука, Москва, 1979
32. Цекова М., А.Георгиева. Оптимизиране процеса на оценяване на риска при строителство на пътища в планински райони. XI МНК по Архитектура и строителство ArCivE 2023, 03.06.2023, Варна, ISBN2535-0781
33. Fisher.R.A., The use of multiple measurements in taxonomic problems. Ann Egenics 1936.II. №7, p.178-179
34. Fisher.R.A., E.Yates, Statistical tables from biological agricultural and medical research. Oliver and Boyd. Edinburgh. 1938.
35. Fucunaga K. Introductions to statistical pattern recognition. 2nd Edition. Academic press, New York 1990

36. Duda R., P.Hort. Pattern Recognition and scene analysis. J. Wiley, New York 1973
37. Karastatev D., D.Radev, N.Kaponov, A.Atanasov. Late quantitative functional prognosis in rechililitated patients with acute myocardial in fraction. IX World Congress of Cardiology, Moscow 1982
38. Nedev A., H. Kucharzewski, a.a. Ein Expertsystem fur die Diagnostik von Schiffsdieselmotoren. Warnemunde 1982.