

МАТЕМАТИКО-СТАТИСТИЧЕСКИЙ АНАЛИЗ ТРАВМАТИЗМА – ПОСТАНОВКА ЗАДАЧИ

For rational planning the facilities on preventive maintenance of the traumatism necessary to define the most high level of his factors on different sign.

By comparing average factors of the traumatism for determined period or comparison of the average factors at different months, numbers of the months, weekdays, watch of the change, days biological rhythms, in different change, on sign flap, beside different enterprise, on type of the work etc. most often does not manage to get the unambiguous answer about value of their difference.

For comparison of the factors of the production traumatism is offered use the method of the checking the statistical hypotheses, which allows with high degree of reliability to get the reliable results about value difference average factors of the production traumatism on different sign.

Введение. Доказано, что несчастный случай – это сложное явление, вызванное рядом причин и факторов. С целью наиболее эффективного использования имеющихся средств необходимо выявить общие закономерности возникновения несчастных случаев и оценить их количественно. Производственный травматизм представляет собой совокупность несчастных случаев. Но в большинстве случаев на современном этапе борьба с травматизмом сводится к снижению его уровня, поэтому важно не только установить причины несчастных случаев, но и определить факторы, способствовавшие их возникновению.

Основная часть. При анализе травматизма возникают следующие задачи:

определение значимости изменения показателя травматизма во времени;

оценка отклонения числа несчастных случаев от их среднего значения (в различные месяцы, числа месяцев, дни недели, часы смены, дни биологических ритмов);

сравнение двух средних показателей (в различные смены, по признаку пола);

сравнение нескольких средних показателей (у различных предприятий, по видам работ и т. п.).

Подобные задачи могут быть успешно решены методами математической статистики.

Наиболее распространенной характеристикой травматизма является коэффициент частоты, выражающий количество несчастных случаев, приходящихся на 1000 работающих:

$$K_{\text{ч}} = \frac{T \times 1000}{P}, \quad (1)$$

где T – общее количество пострадавших за определенный период времени; P – среднесписочная численность работающих за тот же период времени.

Этот показатель позволяет оценить, как часто происходят несчастные случаи, но не несет информации об их последствиях.

Коэффициент тяжести травматизма, выражающий число дней нетрудоспособности, приходящийся на одну травму за исследуемый период времени,

$$K_{\text{т}} = \frac{D}{T}, \quad (2)$$

где D – число дней нетрудоспособности, вызванной несчастными случаями.

В этом случае известна средняя продолжительность потери трудоспособности в результате одной травмы, однако неизвестно, как часто они происходят.

На наш взгляд, наиболее пригоден для статистической обработки коэффициент частоты, позволяющий дать сравнительную оценку частоты травматизма. В некоторых случаях, когда по сравниваемым вариантам количество работающих одинаково, целесообразно использовать непосредственно число несчастных случаев.

При изучении производственного травматизма исследователь получает набор чисел – коэффициентов частоты или количества происшедших несчастных случаев. Эти значения, определяемые большим числом различных факторов, заранее предсказать невозможно, поэтому они называются случайными величинами. Поскольку число несчастных случаев может быть только целым, то мы имеем дело с дискретными величинами.

Коэффициент частоты и число несчастных случаев представляют собой дискретную случайную величину, для анализа которой могут быть применены соответствующие методы математической статистики [1].

Факторы, определяющие уровень травматизма, будем называть дескрипторами. Это, например, возраст, годы, причины и т. д. Каждый дескриптор подразделяется на признаки. Например, 25...30 лет, 2006 год, нарушение правил охраны труда и т. д. Как видим, дескрипторы и признаки могут быть как качественными, так и количественными.

Таким образом, задача анализа травматизма сводится к оценке различия его показателей по различным признакам. При этом пользуются рядом коэффициентов частоты или чисел несчастных случаев, распределенных по различным признакам. Такой ряд называется выбо-

рочной совокупностью или выборкой, в отличие от генеральной совокупности, которая предполагает наличие всех потенциально возможных наблюдений, в то время как травматизм всегда анализируется за конкретный период. Отдельные члены выборки называются вариантами, в дальнейшем будем их обозначать x_i . А число членов, входящих в выборку, обозначается символом n и носит название объема выборки.

Важнейшей характеристикой выборки является среднее арифметическое или, сокращенно, среднее

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i. \quad (3)$$

Перечисленные выше задачи анализа травматизма решаются методами математической статистики, основанными на проверке так называемых статистических гипотез. При этом используется следующая логическая схема. Вначале выдвигается определенная статистическая гипотеза, которая называется нулевой, например: ряд коэффициентов частоты травматизма за последние годы существенно не изменяется или значения коэффициентов частоты у различных предприятий за последние годы существенного различия не имеют. Противоположная гипотеза, отвергающая сделанное предположение, носит название альтернативной или альтернативы. Затем нулевая гипотеза сопоставляется с имеющимся эмпирическим материалом. Если имеющиеся данные не противоречат нулевой гипотезе, то она принимается, в противном случае отвергается в пользу альтернативной.

Для сравнения используются специальные критерии. Принцип их применения основан на вероятностной оценке наблюдающихся отклонений имеющихся данных от нулевой гипотезы. Если эти отклонения невелики и имеют высокую вероятность, то их наличие можно объяснить влиянием различных случайных неучтенных факторов и сами отклонения следует рассматривать как случайные. Очевидно, что основанием для отказа от нулевой гипотезы такие отклонения служить не могут. Наоборот, если отклонения велики, то их нельзя объяснить влиянием случайных факторов. Остается предположить, что причина их появления кроется в несоответствии эмпирического материала выбранной гипотезе. Естественно, что установить четкую границу по вероятности между этими двумя выводами довольно трудно, она всегда будет сохранять известную условность.

Изложенная выше проблема решается с помощью так называемого принципа практической невозможности маловероятных событий. При этом предполагается, что события, имеющие очень низкую вероятность (т. е. почти не-

вероятные), наблюдаться не должны и во внимание могут не приниматься.

События, величины или расхождения в величинах, которые по причине их малой вероятности нельзя считать случайными, именуется в математической статистике значимыми, а наибольшее значение вероятности, при котором событие можно считать значимым (или соответственно наименьшая вероятность, при которой событие можно считать случайным), носит название уровня значимости и обозначается α .

Таким образом, значимым можно считать любое событие, если его вероятность меньше, чем принятый уровень значимости. При исследовании производственного травматизма принимается $\alpha = 0,05$. В этом случае соответствующая ему доверительная вероятность $p = 1 - \alpha = 0,95$.

Специальные критерии, используемые при проверке нулевой гипотезы, представляют собой математические функции, обычно подчиняющиеся какому-нибудь хорошо изученному закону распределения.

По уровню значимости и доверительной вероятности определяют доверительные пределы, называемые также критическими значениями гипотезы. Область между доверительными пределами, где нулевая гипотеза оправдывается, называется областью нулевой гипотезы, а области, расположенные за доверительными пределами, где оправдывается альтернативная гипотеза, носят название критических областей. Таких областей может быть две: они отвечают неравенствам $x < x_{\alpha/2}$ (область неправдоподобно малых значений) и $x > x_{1-\alpha/2}$ (область неправдоподобно больших значений). В этом случае говорят о двусторонней гипотезе или двустороннем критерии.

Если одно из неравенств заведомо невозможно (например, при сравнении двух средних), устанавливают только одно критическое значение x_{α} или $x_{1-\alpha}$. Соответственно будет иметь место только одно неравенство $x < x_{\alpha}$ или $x > x_{1-\alpha}$.

Следует отметить, что при определении критических значений учитывается не объем выборки, а число степеней свободы ν , представляющее собой разность между объемом выборки и числом параметров, по которым выполняется проверка.

При использовании методов проверки статистических гипотез возможны ошибки первого и второго рода. Ошибка первого рода заключается в том, что отвергается нулевая гипотеза, хотя она в действительности верна. Вероятность ошибки первого рода не может быть выше уровня значимости, т. е. всегда остается сравнительно небольшой.

Ошибка второго рода заключается в том, что нулевая гипотеза принимается, хотя она и

неверна. Величина ошибки второго рода обозначается символом β . Способность критерия отвергать нулевую гипотезу, если она ложна, называется силой или мощностью критерия и равна $(1 - \beta)$. Мощность критерия существенно зависит от объема выборки. Казалось бы, при исследовании травматизма увеличение объема выборки не представляет собой серьезной проблемы, так как всегда можно взять коэффициенты частоты или количество несчастных случаев за меньший отрезок времени. Однако к этому следует подходить с большой осторожностью, потому что, как правило, исследователь имеет дело с ограниченным объемом материала, и увеличение числа признаков может существенно увеличить разброс показателей.

Очевидно, что при анализе травматизма, когда проверяются гипотезы о равенстве средних коэффициентов частоты травматизма, ошибка первого рода чревата большими неблагоприятными последствиями по сравнению с ошибкой второго рода. Ошибочное отклонение верной гипотезы о равенстве коэффициентов травматизма приведет к более нерациональному распределению средств на охрану труда, чем в случае принятия ложной гипотезы. Поэтому, на наш взгляд, достаточно задаваться при анализе травматизма значением $\beta = 0,1$.

При проверке статистических гипотез используются две группы критериев: параметрические и непараметрические. Параметрические критерии основаны на сопоставлении параметров распределений, в частности интересующих нас средних. Большинство из них рассчитано на нормальное распределение или на такие распределения, которые могут быть приведены к нормальному после соответствующих преобразований. В этой связи применение этих критериев требует в большинстве случаев предварительной проверки нормальности эмпирических распределений.

Непараметрические критерии могут применяться при любых видах распределений, что значительно расширяет область их использования. Но параметрические критерии являются

более сильными, чем непараметрические.

Таким образом, изложенное выше позволяет установить следующий порядок решения задач, возникающих при анализе травматизма методом проверки гипотез:

формулируются гипотезы: нулевая и альтернативная;

выбирается критерий: параметрический или непараметрический;

в случае использования параметрического критерия проверяется соответствие распределения эмпирических данных нормальному закону. В случае несоответствия рассматривается возможность применения непараметрического критерия;

при использовании параметрического критерия по заданным значениям $\alpha = 0,05$ и $\beta = 0,1$ определяется минимально необходимый объем выборки. Если он больше имеющегося в распоряжении материала, то следует рассмотреть вопрос либо об увеличении выборки, либо о применении другого критерия;

вычисляется выборочная статистика и сравнивается с табличным значением;

принимается нулевая или альтернативная гипотеза. Делается вывод о равенстве средних показателей травматизма либо об их существенном различии.

Заключение. Целью анализа травматизма является установление всех факторов, влияющих на его уровень. Показатели производственного травматизма при простом сравнении их средних значений за определенный период времени чаще всего не дают однозначного ответа на вопрос о значимости их отличия. Задача выявления значимости отличия средних показателей травматизма по каким-либо признакам успешно может быть решена методами математико-статистического анализа.

Литература

1. Джонсон, Н. Статистика и планирование эксперимента в технике и науке. Методы обработки данных / Н. Джонсон, Ф. Лион. – М.: Мир, 1980. – 610 с.