В.Ф. Антонов, Е.К. Козлова,
А.В. Коржуев, А.М. Черныш

ФИЗИКА И БИОФИЗИКА

РУКОВОДСТВО К ПРАКТИЧЕСКИМ ЗАНЯТИЯМ

УЧЕБНОЕ ПОСОБИЕ

ВТОРОЕ ИЗДАНИЕ,
ИСПРАВЛЕННОЕ И ДОПОЛНЕННОЕ

Министерство образования и науки РФ

Рекомендовано ГБОУ ДПО «Российская медицинская
академия последипломного образования»
в качестве учебного пособия
для студентов медицинских специальностей

Регистрационный номер рецензии 37 от 6 марта 2013 года
ФГАУ «Федеральный институт развития образования»

Москва
ИЗДАТЕЛЬСКАЯ ГРУППА
«ГЭОТАР-Медиа»
2013

1. Математические методы обработки результатов экспериментов

ВВЕДЕНИЕ

Знание статистических методов и умение их применять необходимы не только для понимания медико-биологических научных дисциплин, но и также для эффективной работы в любой из областей здравоохранения. Такое знание необходимо для понимания и интерпретации биологических, клинических и лабораторных данных ввиду их вариабельности.

Работник здравоохранения должен уметь интерпретировать результаты лабораторных тестов, клинические наблюдения и измерения, учитывая случайные колебания значений физиологических параметров, возможность ошибки наблюдения и разброс показаний приборов.

В здравоохранении и клинической медицине часто используются различные статистические концепции при принятии решений по таким вопросам, как клинический диагноз, прогнозирование течения заболевания у отдельного больного, выбор для него лечения и т.п.

В настоящее время очевидна необходимость широкого применения статистики в эпидемиологии и организации здравоохранения, поскольку эти области знания имеют дело с сообществами и популяциями, к которым явно приложимы законы статистики.

Знание статистики стало важным для понимания и критической оценки сообщений в научных, в том числе медицинских, журналах.

Можно выделить следующие области применения статистических методов: сбор данных, представление данных, анализ результатов.

В большинстве случаев полезная информация скрыта в массе необработанных данных. Собранные данные надо организовать так, чтобы получить возможность ясно видеть содержащуюся в них информацию. Статистика изучает научные методы сбора, представления и анализа экспериментальных данных.

1.1. ОЦЕНКА ТОЧЕЧНЫХ ХАРАКТЕРИСТИК НОРМАЛЬНОГО РАСПРЕДЕЛЕНИЯ

Цели работы

1. Изучить понятия «генеральная совокупность» и «выборка».
2. Научиться вычислять выборочную среднюю, исправленную выборочную дисперсию, исправленное среднее квадратическое отклонение.
3. Научиться вычислять доверительный интервал для оценки математического ожидания, соответствующий заданной доверительной вероятности.

Литература

1. *Антонов В.Ф. и др.* Физика и биофизика: учебник. — 2-е изд., испр. и доп. — М.: ГЭОТАР-Медиа, 2013.

2. *Морозов Ю.В.* Основы высшей математики и статистики: учебник. — М.: Медицина, 2010.

3. Конспект лекций и данное пособие.

Вопросы теории

Когда мы распространяем представления о конечной группе лиц на другие группы или на всю генеральную совокупность, мы пользуемся информацией о выборке. Когда врач хочет получить представление о составе и состоянии крови пациента, он проводит анализ небольшой выборки крови. Любое значение искомого параметра, вычисленное на основе ограниченного числа опытов, всегда будет содержать элемент случайности. Работники здравоохранения постоянно имеют дело с информацией, базирующейся на ограниченных выборках. Поэтому они должны хорошо представлять себе границы надежности анализа информации на основе выборочных данных.

В биологической и медицинской статистике часто приходится исследовать распределение того или иного признака для весьма большой совокупности индивидуумов, образующих статистический коллектив (таким признаком может быть, например, содержание белка в зерне пшеницы, вес новорожденного ребенка, период колебаний маятника и т.д.). Данный признак является случайной величиной, значение которой от индивидуума к индивидууму меняется. Однако, для того чтобы составить представление о распределении этой случайной величины или о ее важнейших характеристиках, нет необходимости обследовать каждый объект данной обширной (генеральной) совокупности, а можно обследовать некоторую выборку достаточно большого объема для того, чтобы в ней были выявлены существенные черты изучаемого распределения.

Статистическая совокупность представляет собой множество объектов, однородных относительно признака, характеризующего эти объекты.

Генеральной совокупностью называется совокупность, состоящая из всех объектов, которые могут быть охарактеризованы некоторой величиной X. Теоретически это бесконечно большая или приближающаяся к бесконечности совокупность. Число объектов генеральной совокупности называют ее объемом и обозначают N.

Выборочной совокупностью, или *выборкой*, называется множество объектов, случайно отобранных из генеральной совокупности. Число объектов выборки называют ее объемом и обозначают n.

Для того чтобы свойства выборки достаточно хорошо отражали свойства генеральной совокупности, выборка должна быть осуществлена случайно, т.е. все объекты должны иметь одинаковую вероятность попасть в выборку.

Поскольку на практике приходится иметь дело с ограниченным количеством экспериментальных данных, то результаты наблюдений и их обработки содержат больший или меньший элемент случайности.

Характеристики статистического распределения выборки применяются для оценки неизвестных параметров теоретического распределения вероятностей.

Различают точечные оценки случайной величины (одним числом) и интервальные (оценивание параметра совокупности в виде интервала).

Введем некоторые понятия.

Генеральная средняя $\bar{X}_г$ — среднее арифметическое значение признака $X_1, X_2, ..., X_N$ генеральной совокупности, т.е.

$$\bar{X}_г = \frac{1}{N}\sum_{i=1}^{N} X_i.$$

Генеральная средняя равна математическому ожиданию случайной величины:

$$\bar{X}_г = \mu.$$

Выборочная средняя $\bar{X}_в$ — среднее арифметическое значение признака выборочной совокупности $X_1, X_2, ..., X_n$, т.е.

$$\bar{X}_в = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Генеральная дисперсия:

$$D(x) = \sigma^2 = \frac{1}{N}\sum_{i=1}^{N}\left(X_i - \mu\right)^2.$$

Выборочная дисперсия:

$$\sigma_{\text{в}}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_{\text{в}}\right)^2.$$

Точечные оценки. За оценку неизвестного значения μ измеряемой величины принимается выборочная средняя:

$$\mu = \bar{X}_{\text{в}} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

За оценку дисперсии D принимается значение исправленной выборочной дисперсии S^2:

$$D = S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_{\text{в}})^2 = \frac{n}{n-1}\sigma_{\text{в}}^2.$$

Интервальная оценка математического ожидания (доверительный интервал для математического ожидания случайной величины, распределенной по нормальному закону, при неизвестном значении параметра σ)

Пусть случайная величина X имеет нормальное распределение, причем неизвестны параметры μ и σ.

В ряде задач требуется не только найти для параметра μ подходящее численное значение, но и оценить его точность. Требуется знать, к каким ошибкам может привести замена параметра μ его точечной оценкой $\bar{X}_{\text{в}}$ и с какой степенью уверенности можно ожидать, что эти ошибки не выйдут за известные пределы.

Такого рода задачи особенно актуальны при малом числе наблюдений, когда точечная оценка в значительной мере случайна и приближенная замена может привести к серьезным ошибкам.

Чтобы дать представление о точности и надежности в математической статистике, пользуются так называемыми доверительным интервалом и доверительной вероятностью.

Разные выборки дадут разные оценки. Пусть для параметра μ получена из некоторого опыта точечная оценка $\bar{X}_{\text{в}}$. При этом, заменяя μ на $\bar{X}_{\text{в}}$, мы совершаем некоторую ошибку.