Осовский И.А, softmasterhl.awardspace.info,
учёный, «специалист по внедрению ПСПО в системе образования» 2011г.,
физик, математик, программист, эксперт по компьютерам и свободному программному обеспечению.

Оценка статистических данных.

Редкая педагог знакома даже с основами математической статистики, тем не менее, умеющая читать и понимать смысл слов на кнопочках компьютерных программ...,–уверенно кнопкой мышки нажимает на кнопочки для построения диаграмм, статистических расчётов.

Ныне многие работники вовлечены в составление отчётов, имеющих статистические данные. Программное обеспечение адаптированное на уровень знаний пользователей предоставляет им кажущуюся легкость представления результатов в виде таблиц, диаграмм, графиков и т.п. Однако, подготовка данных требует от пользователей минимальной «математической статистической грамотности», которая, как правило, отсутствует у не специалистов в областях точных наук. Кроме того, необходимо различать фактические данные (например, таблицы, диаграммы...) от статистических.
Например, обычно, расчеты на калькуляторах, электронных таблицах выполняются без учёта статистического характера данных, т.е. без учёта разброса данных определяющего достоверность данных в зависимости от разброса, точности, погрешностей…
Нужно ли это понимание представителям «неточных наук», каждая/каждый может определить для себя сама/сам, но, вопрос в том, если уж создаются отчёты со статистикой, то нужно ли делать его «математически статистически грамотно»?

0. Оценка статистических данных является *приближённой*.
Правильная запись приближённого числа состоит из цифр только в *Верных* разрядах.
Верные цифры в записи числа это цифры, (в строгом смысле) находящиеся в старших разрядах от разряда погрешности, иначе это *Сомнительные* цифры.
Приближённое число записывается с одной *сомнительной* цифрой в разряде погрешности указывается с ±погрешностью. Естественно определять погрешность с одной *значащей* цифрой.

1. *Среднее значение* (математическое ожидание) $\bar{x} = \Sigma x_i/n$ – используется для оценки n данных $x_i = \bar{x}+\delta_i$, имеющих статистический, «случайный» характер ($\Sigma\delta_i = 0$).
Погрешность б (разброс) *среднего значения* $\bar{x}$ определяется *среднеквадратичным отклонением* значений данных от среднего, которое является дисперсией[2] нормального распределения случайных величин $б^2 = \Sigma (x_i - \bar{x})^2 / (n*(n-1))$.
В Итоге статистической оценкой является $\bar{x} \pm б$.
Для нескольких **m** средних значений (сумма, разность) погрешности (абсолютные) складываются $б = б_1 + б_2 + ... + б_m$, как приращения (здесь и далее как дифференциалы).
При произведении в любых степенях $\alpha, \beta \dots \gamma$ нескольких оценок друг на друга складываются относительные погрешности $б/X = |\alpha|*б_1/X_1+ |\beta|*б_2/X_2 + ... + |\gamma|*б_m/X_m$.

2. Анализ роста, тенденций и т.п. стат.данных y_i определяется, так называемым, *трендом* или регрессионной зависимостью. Искомая линейная зависимость $Y=aX+b$ является наиболее простым видом, к которому сводятся логарифмы степенных зависимостей.
Коэффициенты **a**, **b** линейной зависимости $y_i = ax_i+b+\delta_i$, ($\Sigma\delta_i = 0$) , находятся из минимума среднеквадратичного отклонения $\overline{\delta^2}$, преобразовав: $y'_i = y_i-\bar{y}$, $x'_i = x_i-\bar{x}$, $b = \bar{y}-a\bar{x}$;

$$d\,\overline{\delta^2}/da = d/da\,(\Sigma(y'_i-ax'_i)^2)/n = 2/n*\Sigma((y'_i-ax'_i)*x'_i) = 2/n*((\Sigma(y'_i*x'_i)-a*\Sigma x'^2_i) = 0$$

$a = \Sigma y'_i*x'_i/\Sigma x'^2_i$ Погрешности: $\sigma_y^2 = \Sigma (y_i-ax_i-b)^2/(n-2)$ Слободянюк А.И. Метод наименьших квадратов в школьном физическом эксперименте//Фізіка: праблемы. Выкладання. 1995. Вып. 1

$$\Delta a = 2\sqrt{\frac{1}{N-2}\left(\frac{S_y^2}{S_x^2} - a^2\right)}, \qquad \Delta b = \sqrt{S_x^2 + \langle x\rangle^2}\,\Delta a. \qquad S_x^2 = \frac{\sum_{i=1}^{N} x_i^2}{N} - \left(\frac{\sum_{i=1}^{N} x_i}{N}\right)^2 = \langle x^2\rangle - \langle x\rangle^2 \qquad S_y^2 = \langle y^2\rangle - \langle y\rangle^2$$