Операции над случайными величинами

Лекция №3

Статистические методы в ядерном эксперименте, ИЯФ, 2024 г.

Описание зависимых случайных величин (I)

Пусть x и y – две случайные величины. Как мы уже видели (см. параграф про свойства дисперсии), в общем случае:

$$D(x+y) = D(x)+D(y)+2(E(xy)-E(x)E(y)) = \sigma_x^2+\sigma_y^2+2(\overline{xy}-\overline{x}\,\overline{y})$$

для независимых случайных величин $E(xy) = E(x)E(y)$, и тогда $D(x+y) = D(x) + D(y)$.

Ковариацией $\mathrm{cov}(x, y)$ двух случайных величин x и y называется число (смешанный второй центральный момент):

$$\mathrm{cov}(x, y) = E((x - E(x))(y - E(y))) = \overline{(x - \overline{x})(y - \overline{y})} = \overline{xy} - \overline{x}\,\overline{y}$$

Таким образом, если $\mathrm{cov}(x, y) \neq 0$, то случайные величины x и y зависимы. Часто бывает проще найти ковариацию двух случайных величин, чем их совместное распределение. Более того, вывод о зависимости случайных величин иногда можно сделать даже не зная их распределений.

Пример: Пусть x и y две независимые случайные величины. Будут ли зависимы x и $z = x + y$?

$$\mathrm{cov}(x, x + y) = E(x(x + y)) - E(x)E(x + y) = D(x) + \mathrm{cov}(x, y) = D(x)$$

Если $D(x) \neq 0$ ($x \neq \mathrm{const}$), то x и $x + y$ зависимы.

Описание зависимых случайных величин (II)

В случае суммы более чем двух случайных величин
$y = a_1x_1 + a_2x_2 + ... + a_nx_n$, получаем:

$$D(y) = \sum_{i=1}^{n} a_i^2 D(x_i) + \sum_{i\neq j} \mathrm{cov}(x_i,\, x_j)a_ia_j = \sum_{i,j=1}^{n} D_{ij}a_ia_j = a^T Da.$$

Симметричная матрица $D_{ij} = D(x_i, x_j) = \mathrm{cov}(x_i,\, x_j)$ называется ковариационной (дисперсионной) матрицей. В экспериментальной физике матрицу D_{ij} часто называют матрицей ошибок.

Однако ковариация все же не очень удобная характеристика зависимости случайных величин:

- Ковариация - размерная величина: $[\mathrm{cov}(x, y)] = [x][y]$.
- Если x и y зависимы, то естественно ожидать, что $x_1 = 10x$ и $y_1 = 10y$ зависимы в той же мере, что и x с y, однако $\mathrm{cov}(x_1, y_1) = 100\,\mathrm{cov}(x, y)$, что делает абсолютную величину ковариации не очень-то удобной характеристикой.
- Хорошо бы отнормировать ковариацию так, чтобы она служила мерой "силы" зависимости случайных величин. Для этого вводят коэффициент корреляции: $\rho(x, y) = \frac{\mathrm{cov}(x,y)}{\sqrt{D(x)}\sqrt{D(y)}} = \frac{\overline{xy} - \overline{x}\,\overline{y}}{\sigma_x\sigma_y}$

Описание зависимых случайных величин (III)

Свойства коэффициента корреляции:

- Если x и y независимы, то $\rho(x, y) = 0$ (обратное неверно!)
- $-1 \leq \rho(x, y) \leq 1$
- $|\rho(x, y)| = 1$ тогда и только тогда, когда x и y связаны линейной зависимостью $y = kx + b$ (k, b – константы)

Замечание 1: Пусть x распределено симметрично относительно 0 с плотностью вероятности $f(x)$, а $y = x^2$. Случайные величины x и y функционально связаны, а значит зависимы, а корреляция $\rho(x, y) = 0$.

Замечание 2: Введём $x' = (x - \overline{x})/\sigma_x$ и $y' = (y - \overline{y})/\sigma_y$, очевидно, что $\overline{x'}, \overline{y'} = 0$, $D(x'), D(y') = 1$, значит $\rho(x, y) = \rho(x', y') = \mathrm{cov}(x', y')$. Вычислим: $D(x' \pm y') = \overline{(x' \pm y')^2} - (\overline{x'} \pm \overline{y'})^2 = D(x') + D(y') \pm 2\mathrm{cov}(x', y') = 2(1 \pm \rho(x', y')) \geq 0 \Rightarrow -1 \leq \rho(x, y) \leq 1$

Замечание 3: Вычислим: $\mathrm{cov}(x, y) = \overline{x(kx + b)} - \overline{x}(k\overline{x} + b) = k\overline{x^2} + b\overline{x} - k\overline{x}^2 - b\overline{x} = kD(x)$, $D(y) = D(kx + b) = k^2 D(x)$, $\rho(x, y) = \mathrm{cov}(x, y)/\sqrt{D(x)D(y)} = k/|k| = \pm 1$

Формула переноса ошибок (I)

Получим приближённую формулу для дисперсии функции $f(x_1, ..., x_n)$ от случайных переменных $\vec{x} = \{x_1, ..., x_n\}$, причём $\overline{x_k} = \mu_k, \; k = 1 \div n$. Разложим $f(\vec{x})$ в ряд Тейлора в окрестности точки $\vec{\mu} = \{\mu_1, ..., \mu_n\}$ ($\frac{\partial f}{\partial x_k}|_{x=\mu} \equiv \frac{\partial f}{\partial \mu_k}$, $\Delta x_k \equiv x_k - \mu_k$):

$$f(\vec{x}) = f(\vec{\mu}) + \sum_{k=1}^{n} \Delta x_k \frac{\partial f}{\partial \mu_k} + \mathcal{O}(\Delta \vec{x}^2),$$

тогда в этом линейном приближении $\overline{f(\vec{x})} = f(\vec{\mu})$, а дисперсия:

$$D(f) = \overline{(f(\vec{x}) - f(\vec{\mu}))^2} = \overline{\left(\sum_{i=1}^{n} \Delta x_i \frac{\partial f}{\partial \mu_i}\right)^2} = \sum_{i,j=1}^{n} \frac{\partial f}{\partial \mu_i} \frac{\partial f}{\partial \mu_j} \overline{\Delta x_i \Delta x_j} = v^T D v,$$

где $\vec{v} = \frac{\partial f}{\partial \vec{\mu}}$ – вектор производных, D – ковариационная матрица.
Если $x_1, ..., x_n$ независимы и $\sigma_{x_k} \ll \overline{x_k}$, то получим формулу переноса ошибок:

$$D(f(x)) = \sum_{k=1}^{n} \left(\frac{\partial f}{\partial \mu_k}\right)^2 D(x_k), \quad \sigma_f = \sqrt{\sum_{k=1}^{n} \left(\frac{\partial f}{\partial \mu_k}\right)^2 \sigma_{x_k}^2}$$

Можно получить поправки следующего порядка к формуле переноса ошибок.

Формула переноса ошибок (II)

Пример 1: В эксперименте измерен квадрат массы τ-лептона $m^2 = (3.158 \pm 0.007)$ ГэВ$^2/c^4$, чему равна масса τ-лептона и её ошибка. Нужно ли учитывать поправку следующего порядка к величине средней массы τ-лептона.

Пример 2: Пусть x и y случайные величины, со средними $\overline{x}$ и $\overline{y}$, и погрешностями (стандартными отклонениями) σ_x и σ_y.
Вычислить ошибку функции $f = x \pm y$, $f = xy$, $f = x/y$, $f = x^n$.

Преобразование случайных величин (I)

Есть случайная величина x с ф.п.в. $f(x)$, ставится задача отыскания ф.п.в $g(h)$ для случайной величины $h = h(x)$.

I) Величина x дискретна.

1) Если существует лишь одно значение $x = x_0$, для которого $h(x) = h(x_0) = h_0$, то $P(h(x) = h_0) = P(h(x) = h(x_0)) = P(x = x_0)$.

2) Если существует несколько значений $x = x_0, x_1, ..., x_m$, для которых $h(x) = h(x_0) = h_0$, то

$$P(h(x) = h_0) = P(x = x_0 \text{ или } x = x_1 \ldots \text{ или } x = x_m) = \sum_{i=0}^{m} P(x = x_i).$$

II) Величина x непрерывна.

1) Если $h(x)$ является однозначной функцией x, то, найдя обратную функцию $x = x(h)$, запишем:

$$f(x)dx = f(x(h))\left|\frac{dx(h)}{dh}\right|dh = g(h)dh, \text{ поэтому } g(h) = f(x(h))\left|\frac{dx(h)}{dh}\right|.$$

2) Если $h(x)$ является неоднозначной функцией x, такой что $h(x_1) = h(x_2) = ... = h(x_m)$, тогда

$P(h \div h + dh) = P(x_1 \div x_1 + dx \text{ или } x_2 \div x_2 + dx \text{ или} \ldots x_m \div x_m + dx) =$

$\sum_{i=1}^{m} P(x_i \div x_i + dx)$ в результате $g(h) = \sum_{h(x_i)=h} f(x_i(h))\left|\frac{dx_i(h)}{dh}\right|$.

Преобразование случайных величин (II)

Можно обобщить преобразование распределения вероятности на случай многомерных распределений. Пусть $(x_1, x_2, \dots x_n)$ нужно преобразовать в $(h_1, h_2, \dots h_n)$, где $h_i = h_i(x_1, x_2, \dots x_n)$, $i = 1, \dots, n$. Плотность вероятности для $\vec{h}$:

$$g(\vec{h}) = f(x_1(\vec{h}), x_2(\vec{h}), \dots, x_n(\vec{h}))\left|\frac{\partial(x_1, \dots, x_n)}{\partial(h_1, \dots, h_n)}\right|,$$

где $f(x_1, x_2, \dots, x_n)$ – исходная совместная плотность вероятности, $\left|\frac{\partial(x_1, \dots, x_n)}{\partial(h_1, \dots, h_n)}\right|$ – модуль Якобиана преобразования.

Пример 1: Найти распределение суммы двух независимых случайных величин x и y, имеющих ф.п.в. $f(x)$ и $g(y)$.

Пример 2: Найти распределение суммы и частного двух независимых случайных величин распределённых согласно ф.п.в. $\gamma(0, 1)$.

Пример 3: Найти распределение частного двух независимых случайных величин x и y (x/y) распределённых по стандартному нормальному закону $(\mu = 0, \sigma = 1)$.