Rozptyl

- Základní vlastnosti disperze

  Var(konst) = 0

  Var(X+Y) = Var(X) + Var(Y) (nezávislé proměnné)
- Lineární změna jednotek Y = rX + s, například z °C na °F. Jak vypočítám střední hodnotu a rozptyl?

Expectation and variance under change of units. For any random variable X and any real numbers r and s,

$$E[rX + s] = rE[X] + s, \quad \text{and} \quad \mathrm{Var}(rX + s) = r^2 \mathrm{Var}(X).$$

- Pozn.: rozptyl je nezávislý na posunu hustoty pravděpodobnosti na ose *x*, protože Var(X) mi určuje jen šířku rozdělení.

Rozptyl

- Př.: stanovení disperze normálního rozdělení $N(\mu, \sigma^2)$

$$\begin{aligned}\operatorname{Var}(X) &= \mathrm{E}\left[(X - \mathrm{E}[X])^2\right] \\ &= \int_{-\infty}^{\infty} (x-\mu)^2 \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \mathrm{d}x \\ &= \sigma^2 \int_{-\infty}^{\infty} z^2 \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}z^2} \mathrm{d}z.\end{aligned}$$

$$\int_{-\infty}^{\infty} z^2 \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}z^2} \mathrm{d}z = 1.$$

Použili jsme substituci:
$z = (x - \mu)/\sigma$

VARIANCE OF A NORMAL DISTRIBUTION. Let X be an $N(\mu, \sigma^2)$ distributed random variable. Then

$$\operatorname{Var}(X) = \int_{-\infty}^{\infty} (x-\mu)^2 \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \mathrm{d}x = \sigma^2.$$

Transformace X

- Už umíme spočítat E[X] a Var[X], pokud dojde ke změně jednotek náhodné proměnné.
- Víme jak pracovat s výpočtem E[X], pokud X transformujeme nějakou funkcí na novou náhodnou proměnnou.
- Jak se změní rozdělení náhodné proměnné, když ji transformujeme na novou náhodnou proměnnou?
- Obecně můžeme novou náhodnou proměnnou Y vytvořit aplikováním zcela obecné funkce *g* na známou náhodnou proměnnou X, tedy:

$$Y = g(X)$$

Transformace X

- U diskrétní náhodné proměnné je transformace proměnné snadná.
- Na každou hodnotu X aplikujeme podle předpisu funkci $g(X)$ a dostaneme novou proměnnou Y.
- Pak jednoduše Y dosadíme do pravděpodobnostní funkce.
- Funguje to stejně i pro spojitou náhodnou proměnnou?
- Bohužel ne, protože f(a) = 0!!! Musel bych spočítat $f(g^{-1}(y))d(g^{-1}(y))$.
- První musíme stanovit distribuční funkci proměnné Y = g(X) a hustotu pravděpodobnosti spočítat derivací F(Y).

Transformace X

- Př.: úsekové měření rychlosti spočívá na měření času při projetí 1 km vzdálenosti. Když auto jede 60 km/h, pak t_1 = 60 s. Když auto jede 90 km/h, pak t_2 = 40 s.
- Náhodná proměnná T bude modelovat čas náhodně vybraného auta na měřeném úseku. Předpokládejme, že T je popsáno rovnoměrným rozdělením s $F_T(t) = P(T \leq t) = (t-40)/20$.
- Jaké bude rozdělení pravděpodobnosti rychlosti měřených aut? V – náhodná proměnná rychlost auta.
- $V = g(T) = (1\ \text{km})/(T/3600) = 3600/T \rightarrow$ transformační funkce.

Transformace X

- Hledáme $F_V(v) = P(V \leq v) = P(3600/T \leq v) =$ $= P(T \geq 3600/v) = 1 - P(T \leq 3600/v) = 1 –$ $((3600/v) – 40)/20 = 3 – 180/v$ pro rychlosti v rozmezí 60 až 90 km/h.
- Hustota pravděpodobnosti $f_V(v)$ pak bude derivace $F_V(v)$ podle dv $\rightarrow f_V(v) = 180/v^2$.
- Př.: Y = 1/X

$$f_Y(y) = \frac{\mathrm{d}}{\mathrm{d}y} F_Y(y) = \frac{1}{y^2} f_X\left(\frac{1}{y}\right) \quad \text{for } y < 0 \text{ and } y > 0.$$

Transformace X

CHANGE-OF-UNITS TRANSFORMATION. Let X be a continuous random variable with distribution function F_X and probability density function f_X. If we change units to $Y = rX + s$ for real numbers $r > 0$ and s, then

$$F_Y(y) = F_X\left(\frac{y-s}{r}\right) \quad \text{and} \quad f_Y(y) = \frac{1}{r} f_X\left(\frac{y-s}{r}\right).$$

- Př.: normální rozdělení

NORMAL RANDOM VARIABLES UNDER CHANGE OF UNITS. Let X be a random variable with an $N(\mu, \sigma^2)$ distribution. For any $r \neq 0$ and any s, the random variable $rX + s$ has an $N(r\mu + s, r^2\sigma^2)$ distribution.

Transformace X

- Jak musím transformovat náhodnou proměnnou, abych získal N(0,1)?
- Pokud zavedu: r = 1/σ a s = -μ/σ, tak transformovaná náhodná proměnná

$$Z = \frac{1}{\sigma}X + \left(-\frac{\mu}{\sigma}\right) = \frac{X-\mu}{\sigma}$$

  má distribuční funkci N(0, 1).
- Tedy každá pravděpodobnostní distribuce N(μ, σ^2) může být transformována na N(0,1) pokud změníme náhodnou proměnnou X na Z.

Transformace X

- Pokud je transformace náhodné proměnné X na Z lineární – $Z = g(X) = r \cdot X + s$, potom $E[g(X)] = g(E[X])$.
- Pokud je transformace nelineární, je třeba provést výpočet $F(Z)$ a pak $f(Z)$ nebo spočítat $f(g^{-1}(Z))d(g^{-1}(Z))$
- Nicméně pro konvexní funkci platí nerovnost:

Jensen's inequality. Let g be a convex function, and let X be a random variable. Then

$$g(\mathrm{E}[X]) \leq \mathrm{E}[g(X)] \,.$$

- Je možné dokázat, že Jensenova nerovnost je ekvivalentní nerovnosti $\mathrm{Var}(X) \geq 0$ (viz definice Var(X)).

Extrémy distribucí

- Mějme sekvenci náhodných proměnných a zajímá nás, pro kterou náhodnou proměnnou nastává extrém (maximum, minimum).
- Př.: nechť X_1, X_2, X_3, …, X_{365} jsou náhodné proměnné popisující výšku hladiny řeky v každém dni v roce. Povodeň nastane tehdy když X_i > „limit“.
- Jaké je rozdělení pravděpodobnosti, že nastane povodeň v budoucnosti? Hledáme pravděpodobnostní distribuci náhodné proměnné Z:

$$Z = \max\{X_1, X_2, \ldots, X_n\}$$

Extrémy distribucí

- Distribuční funkci určíme za předpokladu, že všechny $X_i < a$ a navíc všechny události $X_i < a$ musí být nezávislé.

$$F_Z(a) = \mathrm{P}(Z \leq a) = \mathrm{P}(\max\{X_1, \ldots, X_n\} \leq a)$$
$$= \mathrm{P}(X_1 \leq a, \ldots, X_n \leq a) = \mathrm{P}(X_1 \leq a) \cdots \mathrm{P}(X_n \leq a)$$

THE DISTRIBUTION OF THE MAXIMUM. Let $X_1, X_2, \ldots, X_n$ be n independent random variables with the same distribution function F, and let $Z = \max\{X_1, X_2, \ldots, X_n\}$. Then

$$F_Z(a) = (F(a))^n.$$

Extrémy distribucí

- Lze stanovit i pravděpodobnostní distribuci minima sekvence náhodných proměnných?

$$V = \min\{X_1, X_2, \ldots, X_n\}$$

- Hledáme takovou distribuci, kdy pro všechna X_i bude platit $X_i > a$ a zároveň tyto události budou nezávislé.
- Použijeme trik: hledáme vlastně V takové, že bude komplementární k Z.

Extrémy distribucí

$$F_V(a) = \mathrm{P}(V \le a) = 1 - \mathrm{P}(V > a) = 1 - \mathrm{P}(\min\{X_1, \ldots, X_n\} > a)$$
$$= 1 - \mathrm{P}(X_1 > a, \ldots, X_n > a)\,.$$

$$F_V(a) = 1 - \mathrm{P}(X_1 > a, \ldots, X_n > a) = 1 - \mathrm{P}(X_1 > a) \cdots \mathrm{P}(X_n > a)$$
$$= 1 - (1 - \mathrm{P}(X_1 \le a)) \cdots (1 - \mathrm{P}(X_n \le a)).$$

THE DISTRIBUTION OF THE MINIMUM. Let $X_1, X_2, \ldots, X_n$ be n independent random variables with the same distribution function F, and let $V = \min\{X_1, X_2, \ldots, X_n\}$. Then

$$F_V(a) = 1 - (1 - F(a))^n.$$

Sdružená distribuce pravděpodobnosti

Sdružené distribuce

- V praxi se často setkáváme se situací, že náhodný experiment má výstup více náhodných proměnných.
- Př.: může být třeba sčítání lidu, kdy máme náhodné proměnné: příjem, věk, vzdělání, pohlaví, rodinný stav atp. Může nás zajímat studovat tyto náhodné proměnné společně v jedné pravděpodobnostní distribuci a tím třeba získat další informace (třeba o míře emancipace žen).
- Umožňuje nám to dát pohled na vzájemnou provázanost jednotlivých náhodných proměnných.

Pravděpodobnostní funkce pro X, Y

- Obecně **sdružená pravděpodobnostní funkce** dvou diskrétních náhodných proměnných X a Y je definována na množině všech možných náhodných jevů Z a daná pravděpodobnostmi všech možných jevů (X, Y).

DEFINITION. The *joint probability mass function* p of two discrete random variables X and Y is the function $p : \mathbb{R}^2 \to [0,1]$, defined by

$$p(a,b) = \mathrm{P}(X = a, Y = b) \quad \text{for } -\infty < a, b < \infty.$$

- Diskrétní pravděpodobnostní funkce pak může být dána výčtem všech možných hodnot $p(a_i, b_j)$.

Pravděpodobnostní funkce pro X, Y

- Př.: hod dvěma kostkami. S – součet čísel na kostkách, M – hozené maximum. Pravděpodobnostní funkce je daná tabulkou.

Joint probability mass function $p(a, b) = \mathrm{P}(S = a, M = b)$.

	b					
a	1	2	3	4	5	6
2	1/36	0	0	0	0	0
3	0	2/36	0	0	0	0
4	0	1/36	2/36	0	0	0
5	0	0	2/36	2/36	0	0
6	0	0	1/36	2/36	2/36	0
7	0	0	0	2/36	2/36	2/36
8	0	0	0	1/36	2/36	2/36
9	0	0	0	0	2/36	2/36
10	0	0	0	0	1/36	2/36
11	0	0	0	0	0	2/36
12	0	0	0	0	0	1/36

Pravděpodobnostní funkce pro X, Y

- Lze získat z p(a, b) jednotlivé pravděpodobnostní funkce p(a) a p(b)?
- Musím najít pravděpodobnostní funkci pro jevy {S=1, M=[1, 6]}, {S=2, M=[1, 6]}, {S=3, M=[1, 6]}, …, {S=12, M=[1, 6]}.
- Např.: $p_S(6)$ je dáno

$$\begin{aligned} p_S(6) = \mathrm{P}(S=6) &= \mathrm{P}(S=6, M=1) + \cdots + \mathrm{P}(S=6, M=6) \\ &= p(6,1) + p(6,2) + \cdots + p(6,6) \\ &= 0 + 0 + \frac{1}{36} + \frac{2}{36} + \frac{2}{36} + 0 \\ &= \frac{5}{36}. \end{aligned}$$

Pravděpodobnostní funkce pro X, Y

- Nakonec všechny 3 pravděpodobnostní funkce dostaneme formou výčtu (tabulky.)

Table 9.2. Joint distribution and marginal distributions of S and M.

	b						
a	1	2	3	4	5	6	$p_S(a)$
2	1/36	0	0	0	0	0	1/36
3	0	2/36	0	0	0	0	2/36
4	0	1/36	2/36	0	0	0	3/36
5	0	0	2/36	2/36	0	0	4/36
6	0	0	1/36	2/36	2/36	0	5/36
7	0	0	0	2/36	2/36	2/36	6/36
8	0	0	0	1/36	2/36	2/36	5/36
9	0	0	0	0	2/36	2/36	4/36
10	0	0	0	0	1/36	2/36	3/36
11	0	0	0	0	0	2/36	2/36
12	0	0	0	0	0	1/36	1/36
$p_M(b)$	1/36	3/36	5/36	7/36	9/36	11/36	1

Pravděpodobnostní funkce pro X, Y

- Pravděpodobnostní funkce jedné náhodné proměnné získaná ze sdružené pravděpodobnostní funkce se nazývá jako tzv. **marginální pravděpodobnostní funkce**.
- Je nutné si uvědomit, že sdružená pravděpodobnostní funkce mi nese více informací než jednotlivé pravděpodobnostní distribuce.
- Ve většině případů nelze ze znalostí jednotlivých pravděpodobnostních funkcí získat sdruženou pravděpodobnostní funkci (naopak je to snadné) – viz tabulka na předchozí stránce.

Distribuční funkce pro X, Y

DEFINITION. The *joint distribution function* F of two random variables X and Y is the function $F : \mathbb{R}^2 \to [0,1]$ defined by

$$F(a,b) = \mathrm{P}(X \le a, Y \le b) \quad \text{for } -\infty < a, b < \infty.$$

- Opět jako pro pravděpodobnostní funkci dvou náhodných proměnných můžeme získat marginální F_X a F_y ze sdružené F.

FROM JOINT TO MARGINAL DISTRIBUTION FUNCTION. Let F be the joint distribution function of random variables X and Y. Then the marginal distribution function of X is given for each a by

$$F_X(a) = \mathrm{P}(X \le a) = F(a, +\infty) = \lim_{b \to \infty} F(a,b), \qquad (9.1)$$

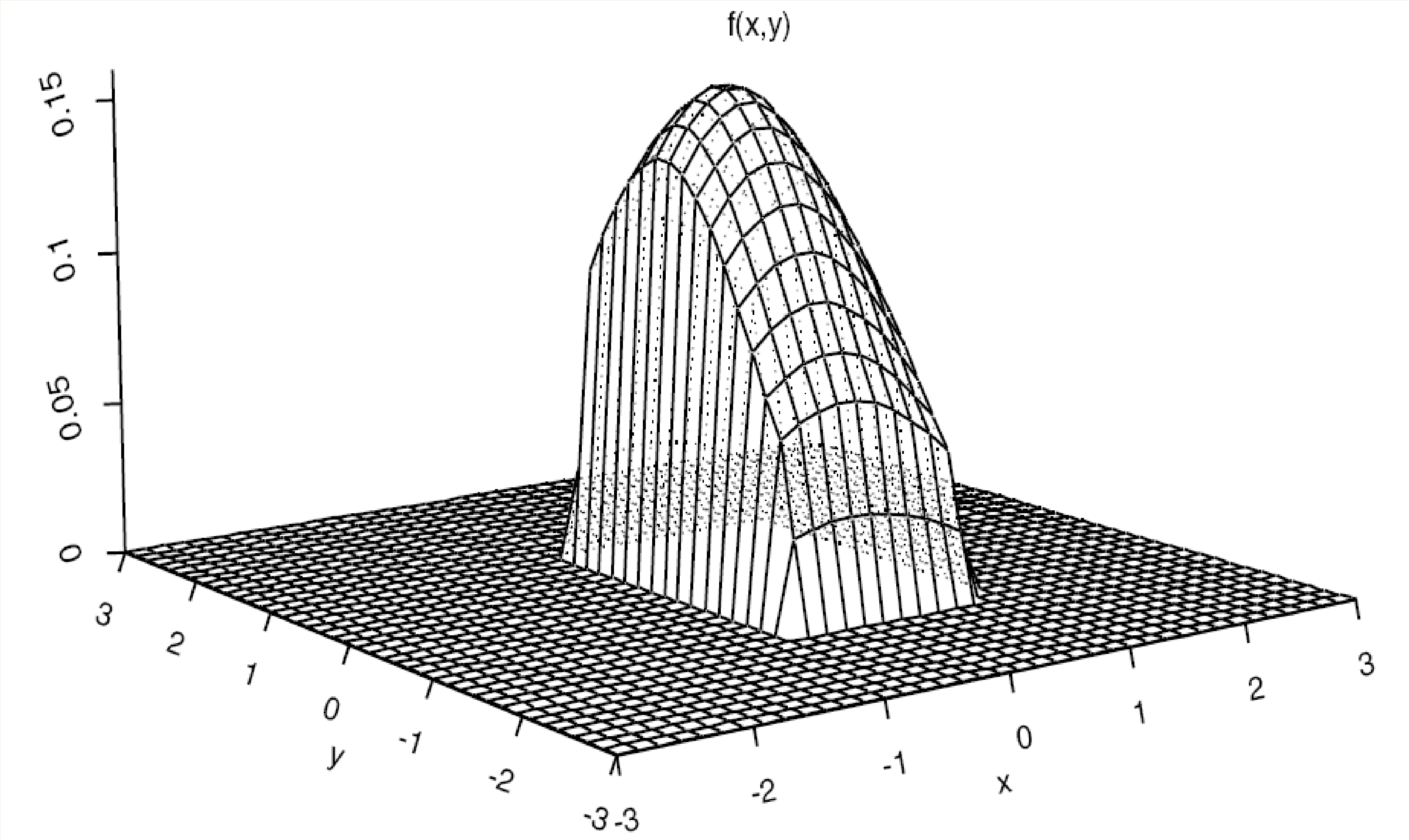
and the marginal distribution function of Y is given for each b by

$$F_Y(b) = \mathrm{P}(Y \le b) = F(+\infty, b) = \lim_{a \to \infty} F(a,b). \qquad (9.2)$$

Spojitá náhodná proměnná

- Budeme hledat hustotu pravděpodobnosti spojitých náhodných proměnných (X, Y).
- Pro jednu spojitou náhodnou proměnnou platí: pravděpodobnost, že X nabývá hodnot v intervalu [a, b] je dána plochou pod křivkou f(x) na intervalu [a, b].
- Pro dvě náhodné proměnné X a Y platí: pravděpodobnost, že (X, Y) nabývá hodnot v „obdélníku“ $[a_1, b_1]\times[a_2, b_2]$ je dána objemem tělesa, které vznikne nad „obdélníkem“ vymezujícím plochu pod křivkou f(x, y).

Spojitá náhodná proměnná



Spojitá náhodná proměnná

DEFINITION. Random variables X and Y have a *joint continuous distribution* if for some function $f : \mathbb{R}^2 \to \mathbb{R}$ and for all numbers a_1, a_2 and b_1, b_2 with $a_1 \le b_1$ and $a_2 \le b_2$,

$$P(a_1 \le X \le b_1, a_2 \le Y \le b_2) = \int_{a_1}^{b_1} \int_{a_2}^{b_2} f(x, y)\, dx\, dy.$$

The function f has to satisfy $f(x, y) \ge 0$ for all x and y, and $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\, dx\, dy = 1$. We call f the *joint probability density function* of X and Y.

- Sdružená distribuční funkce dvou proměnných

$$F(a, b) = \int_{-\infty}^{a} \int_{-\infty}^{b} f(x, y)\, dx\, dy \quad \text{and} \quad f(x, y) = \frac{\partial^2}{\partial x \partial y} F(x, y).$$

Spojitá náhodná proměnná

- Ilustrativní příklad je f(x,y)=0 pro *x* a *y* mimo definiční interval.

$$f(x,y) = \frac{2}{75}\left(2x^2y + xy^2\right) \quad \text{for } 0 \le x \le 3 \text{ and } 1 \le y \le 2.$$

$$\begin{aligned} P\left(1 \le X \le 2, \frac{4}{3} \le Y \le \frac{5}{3}\right) &= \int_1^2 \int_{\frac{4}{3}}^{\frac{5}{3}} f(x,y)\,\mathrm{d}x\,\mathrm{d}y \\ &= \frac{2}{75}\int_1^2 \left(\int_{\frac{4}{3}}^{\frac{5}{3}} (2x^2y + xy^2)\,\mathrm{d}y\right)\mathrm{d}x \\ &= \frac{2}{75}\int_1^2 \left(x^2 + \frac{61}{81}x\right)\mathrm{d}x = \frac{187}{2025}. \end{aligned}$$

Spojitá náhodná proměnná

- Spočítáme si sdruženou distribuční funkci sdružené hustoty pravděpodobnosti:

$$\begin{aligned}
F(a,b) = \mathrm{P}(X \le a, Y \le b) &= \int_{-\infty}^{a} \left(\int_{-\infty}^{b} f(x,y)\,\mathrm{d}y \right) \mathrm{d}x \\
&= \frac{2}{75} \int_{0}^{a} \left(\int_{1}^{b} (2x^2 y + xy^2)\,\mathrm{d}y \right) \mathrm{d}x \\
&= \frac{1}{225} (2a^3 b^2 - 2a^3 + a^2 b^3 - a^2).
\end{aligned}$$

- Pokud parametry *a* a *b* budou mimo definiční interval, pak F(a, b) bude jiná.

Spojitá náhodná proměnná

- Toto zjištění nás pak navede ke stanovení marginální hustoty pravděpodobnosti a distribuční funkce bivariabilní spojité náhodné proměnné.

From joint to marginal probability density function. Let f be the joint probability density function of random variables X and Y. Then the *marginal* probability densities of X and Y can be found as follows:

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\, \mathrm{d}y \quad \text{and} \quad f_Y(y) = \int_{-\infty}^{\infty} f(x, y)\, \mathrm{d}x.$$

- Tedy marginální *f* dostanu integrováním *f* přes tu náhodnou proměnnou, která mě zrovna nezajímá.

n náhodných proměnných

- Obecně lze sdruženou pravděpodobnostní distribuci zobecnit na *n* náhodných proměnných jak pro diskrétní, tak i spojitou náhodnou proměnnou.
- V podstatě je dostačující stanovit sdruženou distribuční funkci:

$$F(a_1, a_2, \ldots, a_n) = \mathrm{P}(X_1 \leq a_1, X_2 \leq a_2, \ldots, X_n \leq a_n)$$

- Sdruženou pravděpodobnostní funkci pak lze stanovit výčtem:

$$p(a_1, a_2, \ldots, a_n) = \mathrm{P}(X_1 = a_1, X_2 = a_2, \ldots, X_n = a_n)$$

- Pro *n* spojitých proměnných pak hustotu pravděpodobnosti získáme parciální derivací sdružené F.

Nezávislost náhodných proměnných

- Intuitivně tušíme, že náhodné proměnné X a Y jsou nezávislé, pokud náhodné jevy týkající se každé náhodné proměnné jsou nezávislé.

DEFINITION. The random variables X and Y, with joint distribution function F, are *independent* if

$$P(X \le a, Y \le b) = P(X \le a)\, P(Y \le b)\,,$$

that is,

$$F(a,b) = F_X(a) F_Y(b) \tag{9.4}$$

for all possible values a and b. Random variables that are not independent are called *dependent*.

Nezávislost náhodných proměnných

- Jinými slovy nezávislost náhodných proměnných nám garantuje, že sdruženou pravděpodobnostní distribuci lze rozložit na marginální distribuce.
- Pokud náhodné proměnné X a Y jsou nezávislé, pak:

$$\mathrm{P}(X \in A, Y \in B) = \mathrm{P}(X \in A)\,\mathrm{P}(Y \in B),$$

kde A a B jsou buď intervaly nebo jednotlivé hodnoty, kterých náhodné proměnné nabývají.

- Parciální derivací rovnice (9.4) podle každé náhodné proměnné dostaneme podmínku nezávislosti pro hustotu pravděpodobnosti: $f(x,y) = f_X(x) f_Y(y)$
- Podmínku nezávislosti pak jednoduše můžeme rozšířit na *n* náhodných proměnných.

Přenášení nezávislosti

- Zachovává se nezávislost dvou náhodných proměnných X a Y pokud provedeme transformaci náhodných proměnných?
- Je možné dokázat, že pokud jsou X_1, X_2, X_3,…, X_n nezávislé, tak i transformované náhodné proměnné Y_1, Y_2, Y_3,…, Y_n jsou také nezávislé.

PROPAGATION OF INDEPENDENCE. Let $X_1, X_2, \ldots, X_n$ be independent random variables. For each i, let $h_i : \mathbb{R} \to \mathbb{R}$ be a function and define the random variable

$$Y_i = h_i(X_i).$$

Then $Y_1, Y_2, \ldots, Y_n$ are also independent.

Střední hodnota sdružené distribuce

- Bude nás zajímat, jak spočítat střední hodnotu nebo rozptyl náhodných proměnných majících sdruženou pravděpodobnostní distribuci.
- Př.: továrna vyrábí válcové výrobky, kdy poloměr R a výška H jsou díky použité technologii v určitém tolerančním pásmu dané rovnoměrným rozdělením náhodné proměnné. Bude nás zajímat jaký bude objem výrobku $V=\pi R^2 H$. V je tedy sdružená náhodná proměnná. Jaká bude velikost E[V]?

Střední hodnota sdružené distribuce

$$\mathrm{E}[V] = \int_{-\infty}^{\infty} v f_V(v)\,\mathrm{d}v$$

- Bohužel neznáme hustotu pravděpodobnosti sdružené distribuce $f_V(v)$.
- Využijeme pravidla pro počítání E[X] pro složenou funkci – viz stránka 38 v přednášce 2:

$$\mathrm{E}[V] = \mathrm{E}\left[\pi H R^2\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \pi h r^2 f(h,r)\,\mathrm{d}h\,\mathrm{d}r$$

- Pokud jsou náhodné proměnné H a R nezávislé, můžeme využít marginálních distribucí pro každou náhodnou proměnnou H a R:

$$\mathrm{E}[V] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \pi h r^2 f_H(h) f_R(r)\,\mathrm{d}h\,\mathrm{d}r$$

Střední hodnota sdružené distribuce

TWO-DIMENSIONAL CHANGE-OF-VARIABLE FORMULA. Let X and Y be random variables, and let $g : \mathbb{R}^2 \to \mathbb{R}$ be a function.
If X and Y are *discrete* random variables with values $a_1, a_2, \ldots$ and $b_1, b_2, \ldots$, respectively, then

$$\mathrm{E}\left[g(X,Y)\right] = \sum_i \sum_j g(a_i, b_j)\mathrm{P}(X = a_i, Y = b_j)\,.$$

If X and Y are *continuous* random variables with joint probability density function f, then

$$\mathrm{E}\left[g(X,Y)\right] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) f(x,y)\,\mathrm{d}x\,\mathrm{d}y.$$

Střední hodnota sdružené distribuce

- Pokud bude transformace náhodných proměnných lineární pak výpočet E[X] se zjednoduší:

LINEARITY OF EXPECTATIONS. For all numbers r, s, and t and random variables X and Y, one has

$$\mathrm{E}[rX + sY + t] = r\mathrm{E}[X] + s\mathrm{E}[Y] + t.$$

- Opět to lze rozšířit na *n* náhodných proměnných.
- Pravidlo lze například použít na výpočet střední hodnoty binomického rozdělení Bin(n, p). Výpočet podle definice E[X] není triviální:

$$\mathrm{E}[X] = \sum_{k=0}^{n} k\mathrm{P}(X = k) = \sum_{k=0}^{n} k\binom{n}{k} p^k (1-p)^{n-k}.$$

- Dá se ukázat, že náhodná proměnná X s binomickým rozdělením Bin(n,p) může být vyjádřena jako $X = R_1 + R_2 + R_3 + \ldots R_n$, kde R_i je náhodná proměnná s Bernoulliho rozdělením Ber(p). Střední hodnota Ber(p) je rovna *p*. Potom na základě linearity výpočtu střední hodnoty dostaneme:

$$\mathrm{E}[X] = \mathrm{E}[R_1] + \mathrm{E}[R_2] + \cdots + \mathrm{E}[R_n] = np.$$

Kovariance

- Z pravidla linearity střední hodnoty pro sdruženou náhodnou proměnnou platí, že E[X+Y] = E[X] + E[Y].
- Platí podobné lineární pravidlo i pro E[XY] nebo pro výpočet rozptylu Var(X+Y)?
- Pozn.: Pro nezávislé náhodné proměnné to platí!!!
- Na straně 25 jsme uváděli příklad sdružené hustoty pravděpodobnosti: $f(x,y) = \frac{2}{75}(2x^2y + xy^2) \quad \text{for } 0 \leq x \leq 3 \text{ and } 1 \leq y \leq 2$
- Snadno nahlédneme, že:

$$\mathrm{Var}(X+Y) = \frac{939}{2000} \quad \text{and} \quad \mathrm{Var}(X) + \mathrm{Var}(Y) = \frac{989}{2500} + \frac{791}{10\,000} = \frac{4747}{10\,000}$$

- Jasně vidíme, že **neplatí** rovnost Var(X+Y) = Var(X)+Var(Y).
- Odvodíme si vztah pro výpočet Var(X+Y):

Kovariance

$$\mathrm{Var}(X+Y) = \mathrm{E}\left[(X+Y-\mathrm{E}[X+Y])^2\right].$$

Now $X+Y-\mathrm{E}[X+Y] = (X-\mathrm{E}[X]) + (Y-\mathrm{E}[Y])$, so that

$$(X+Y-\mathrm{E}[X+Y])^2 = (X-\mathrm{E}[X])^2 + (Y-\mathrm{E}[Y])^2 + 2\,(X-\mathrm{E}[X])\,(Y-\mathrm{E}[Y])$$

- Tedy dostaneme ve finále

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{E}\left[(X-\mathrm{E}[X])(Y-\mathrm{E}[Y])\right]$$

- Tedy rozptyl součtu dvou náhodných proměnných je roven součtu jejich rozptylů plus navíc člen vyjadřující vzájemné ovlivňování jednotlivých náhodných proměnných.

Kovariance

Definition. Let X and Y be two random variables. The *covariance* between X and Y is defined by

$$\mathrm{Cov}(X,Y) = \mathrm{E}\left[(X - \mathrm{E}[X])(Y - \mathrm{E}[Y])\right].$$

- Jestliže Cov(X, Y) je kladná, tak pokud X bude větší jak E[X], tak s velkou pravděpodobností bude Y větší jak E[Y] a naopak. Říkáme, že X a Y jsou **pozitivně korelovány**.
- Jestliže Cov(X, Y) je záporná, tak nastane opačný efekt a říkáme, že X a Y jsou **negativně korelovány**.
- Jestliže Cov(X, Y) = 0, tak X a Y jsou **nekorelovány**.

Kovariance

AN ALTERNATIVE EXPRESSION FOR THE COVARIANCE. Let X and Y be two random variables, then

$$\mathrm{Cov}(X,Y) = \mathrm{E}[XY] - \mathrm{E}[X]\,\mathrm{E}[Y]\,.$$

- Tedy kovariance mi měří určitý druh závislosti mezi náhodnými proměnnými.
- Výše uvedený vztah taky ilustruje skutečnost, že obecně střední hodnota součinu náhodných proměnných není rovna součinu středních hodnot náhodných proměnných

$$E[XY] \neq E[X] \cdot E[Y].$$

Nezávislost vs. nekorelace

- Mějme dvě **nezávislé** náhodné proměnné X a Y, očekáváme, že spolu nebudou nijak souviset, tedy, že budou **nekorelované**.
- Dá se ukázat, že pro diskrétní i spojitou náhodnou proměnnou X a Y, splňující výše uvedenou podmínku bude platit, že: E[XY]=E[X]·E[Y].
- Z toho plyne důležité tvrzení:

INDEPENDENT VERSUS UNCORRELATED. If two random variables X and Y are independent, then X and Y are uncorrelated.

- Opačné tvrzení ale obecně neplatí!!! Tedy nekorelované náhodné proměnné nemusí být nutně nezávislé.

Nezávislost vs. nekorelace

Variance of the sum. Let X and Y be two random variables. Then always

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X,Y)\,.$$

If X and Y are *uncorrelated*,

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)\,.$$

- Vždy platí, že E[X+Y] = E[X] + E[Y]. Ale rovnost Var(X+Y) = Var(X) + Var(Y) platí jen tehdy, když X a Y jsou nekorelované.
- Výše uvedené platí i pro výpočet rozptylu součtu *n* nekorelovaných náhodných proměnných.

Nezávislost vs. nekorelace

- Pravidlo pro výpočet Var(X+Y) lze opět aplikovat na výpočet rozptylu Bin(n,p) podobně jako u výpočtu střední hodnoty Bin(n,p).
- Náhodnou proměnnou X popsanou Bin(n,p) lze rozepsat jako X = R_1+ R_2+ R_3+… R_n, kde R_i je náhodná proměnná s rozdělením Ber(p).
- Platí:

$$\begin{aligned}\mathrm{Var}(R_i) &= \mathrm{E}\left[R_i^2\right] - \left(\mathrm{E}\left[R_i\right]\right)^2 = 0^2 \cdot (1-p) + 1^2 \cdot p - \left(\mathrm{E}\left[R_i\right]\right)^2 \\ &= p - p^2 = p(1-p).\end{aligned}$$

- Za předpokladu, že všechny R_i jsou nezávislé (a tedy nekorelované), pak:

$$\mathrm{Var}(X) = \mathrm{Var}(R_1) + \mathrm{Var}(R_2) + \cdots + \mathrm{Var}(R_n) = np(1-p).$$

Korelační koeficient

- Kovariance mezi náhodnými proměnnými nám indikuje, jak se náhodné proměnné vzájemně ovlivňují.
- Z rovnice pro Cov(X,Y) je zřejmé, že kovariance závisí na jednotkách v jakých jsou jednotlivé náhodné proměnné vyjádřeny. Jinými slovy, při transformaci proměnných na nové jednotky se mi změní i Cov(X,Y), ale vzájemné ovlivnění proměnných se mi logicky změnit nemůže.

COVARIANCE UNDER CHANGE OF UNITS. Let X and Y be two random variables. Then

$$\mathrm{Cov}(rX + s, tY + u) = rt\,\mathrm{Cov}(X, Y)$$

for all numbers r, s, t, and u.

Korelační koeficient

- V mnoha situacích kovariance není vhodná pro vyjádření závislosti dvou náhodných proměnných X a Y.
- Proto se zavádí standardizovaná veličina **korelační koeficient**:

DEFINITION. Let X and Y be two random variables. The *correlation coefficient* $\rho(X,Y)$ is defined to be 0 if $\mathrm{Var}(X)=0$ or $\mathrm{Var}(Y)=0$, and otherwise

$$\rho(X,Y)=\frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\,\mathrm{Var}(Y)}}.$$

- Korelační koeficient je **bezrozměrná veličina** a jeho velikost není ovlivněná změnou proměnných. Jen je možné ovlivnit znaménko korelačního koeficientu.
- Tedy platí:

$$\rho(rX+s,tY+u)=\begin{cases}-\rho(X,Y) & \text{if } rt<0,\\ \rho(X,Y) & \text{if } rt>0.\end{cases}$$

Korelační koeficient

- Dvě náhodné proměnné X a Y jsou nejvíce korelované pokud:
  - X = Y, tedy $\rho(X, Y) = 1$
  - X = -Y, tedy $\rho(X, Y) = -1$
- Tedy pro nekonstantní náhodné proměnné platí: $-1 \leq \rho(X,Y) \leq 1.$
- Důkaz:

$$
\begin{aligned}
0 \leq \operatorname{Var}\left(\frac{X}{\sqrt{\operatorname{Var}(X)}} + \frac{Y}{\sqrt{\operatorname{Var}(Y)}}\right)
&= \operatorname{Var}\left(\frac{X}{\sqrt{\operatorname{Var}(X)}}\right) + \operatorname{Var}\left(\frac{Y}{\sqrt{\operatorname{Var}(Y)}}\right) \\
&\quad + 2\operatorname{Cov}\left(\frac{X}{\sqrt{\operatorname{Var}(X)}}, \frac{Y}{\sqrt{\operatorname{Var}(Y)}}\right) \\
&= \frac{\operatorname{Var}(X)}{\operatorname{Var}(X)} + \frac{\operatorname{Var}(Y)}{\operatorname{Var}(Y)} + \frac{2\operatorname{Cov}(X,Y)}{\sqrt{\operatorname{Var}(X)\operatorname{Var}(Y)}} = 2\left(1 + \rho(X,Y)\right).
\end{aligned}
$$

- Poslední nerovnost implikuje $\rho(X, Y) \geq -1$. Pokud místo X bude –X, pak $\rho(X, Y) \leq 1$.

Korelační koeficient

- V literatuře se ρ(X, Y) také nazývá jako Pearsonův korelační koeficient.
- Tedy korelační koeficient nám umožňuje stanovit sílu vztahu mezi dvěma náhodnými proměnnými.
- Jde tedy o dvoustranný reciproční vztah dvou náhodných proměnných X a Y – jsou tedy vzájemně na sobě závislé. Tedy náhodné prvky jedné proměnné mají tendenci se vyskytovat společně s určitými hodnotami druhé náhodné proměnné.
- Př.: vzájemný vztah mezi délkou předních a zadních končetin, vztah mezi rozpětím křídel a délkou ocasu ptáků atp. Tyto náhodné proměnné jsou vzájemně korelovány.