# Kapitola 2.: Uspořádání pokusů

## 2.1. Motivace

Abychom mohli správně vyhodnotit výsledky pokusu, musí být pokus dobře naplánován. V závislosti na záměrech experimentátora rozeznáváme několik typů uspořádání pokusů: jednoduché pozorování (zkoumají se hodnoty náhodné veličiny pozorované za týchž podmínek), dvojné pozorování (zkoumá se rozdílnost hodnot náhodné veličiny pozorované za dvojích různých podmínek) a mnohonásobné pozorování (zkoumá se rozdílnost hodnot náhodné veličiny pozorované za $r \geq 3$ různých podmínek). Podle typu uspořádání pokusu pak volíme vhodnou statistickou metodu.

V této kapitole probereme pouze ty nejjednodušší typy uspořádání pokusů. V praxi (např. v medicínském nebo zemědělském výzkumu) používají vědci často velmi složité plány experimentů. V doporučené literatuře je plánování experimentů věnována podkapitola 2.4.

V následujícím textu se zaměříme na situaci, kdy zkoumáme hmotnostní přírůstky stejně starých selat téhož plemene při různých výkrmných dietách. Určitou výkrmnou dietu aplikujeme např. po dobu půl roku. Každý den zjišťujeme hmotností přírůstky každého selete a po uplynutí půl roku vypočteme pro každé sele průměrný hmotnostní přírůstek.

## 2.2. Jednoduché pozorování

Náhodná veličina je pozorována za týchž podmínek. Situace je charakterizována jedním náhodným výběrem $X_1, ..., X_n$. (Náhodně vybereme n stejně starých selat téhož plemene, podrobíme je jediné výkrmné dietě a zjistíme hmotnostní přírůstky. Tak dostaneme realizaci jednoho náhodného výběru.)

Pokud lze očekávat, že náhodný výběr pochází z normálního rozložení, můžeme např. konstruovat interval spolehlivosti pro neznámou střední hodnotu, neznámý rozptyl či směrodatnou odchylku průměrných denních hmotnostních přírůstků nebo testovat hypotézu, že střední hodnota průměrných denních hmotnostních přírůstků neklesne pod určitou hranici. (Tyto úkoly budeme řešit ve 4. kapitole.)

## 2.3. Dvojné pozorování

Zkoumá se rozdílnost hodnot náhodné veličiny pozorované za dvojích různých podmínek. Existují dvě odlišná uspořádání tohoto pokusu.

**2.3.1. Dvouvýběrové porovnávání**: Situace je charakterizována dvěma nezávislými náhodnými výběry $X_{11},\ldots,X_{1n_1}$ a $X_{21},\ldots,X_{2n_2}$. (Z populace všech dostupných stejně starých selat téhož plemene náhodně vybereme $n_1 + n_2$ jedinců. Náhodně je rozdělíme na dva soubory o rozsazích $n_1$ a $n_2$, první podrobíme výkrmné dietě č. 1 a druhý výkrmné dietě č. 2. Tak dostaneme realizace dvou nezávislých náhodných výběrů.)

Za předpokladu, že dané náhodné výběry pocházejí z normálních rozložení, lze např. konstruovat interval spolehlivosti pro rozdíl středních hodnot či podíl rozptylů průměrných denních hmotnostních přírůstků nebo testovat hypotézu o stejné účinnosti obou výkrmných diet. (Tyto úkoly budeme řešit v 5. kapitole.)

**2.3.2. Párové porovnávání**: Situace je charakterizována jedním náhodným výběrem $(X_{11},X_{12})$, ..., $(X_{n1},X_{n2})$ z dvourozměrného rozložení. Párem se rozumí dvojice $(X_{i1},X_{i2})$, i = 1, ..., n . Úloha se zpravidla převádí na jednoduché pozorování náhodného výběru rozdílů $X_{i1} - X_{i2}$, kde i = 1, ..., n. (Náhodně vybereme n vrhů stejně starých selat téhož plemene a z nich vždy dva sourozence a náhodně jim přiřadíme 1. a 2. výkrmnou dietu. Tak dostaneme realizaci náhodného výběru z dvourozměrného rozložení.)

Lze-li dvourozměrný náhodný výběr považovat za výběr z dvourozměrného normálního rozložení, budeme se zabývat konstrukcí intervalu spolehlivosti pro rozdíl středních hodnot průměrných denních hmotnostních přírůstků nebo testovat hypotézu o stejné účinnosti obou výkrmných diet. (Řešení úkolů tohoto typu je popsáno ve 4. kapitole.)

## 2.3. Mnohonásobné pozorování

Zkoumá se rozdílnost hodnot náhodné veličiny pozorované za $r \geq 3$ různých podmínek. Existují dvě odlišná uspořádání tohoto pokusu.

**2.3.1. Mnohovýběrové porovnávání**: Situace je charakterizována r nezávislými náhodnými výběry $X_{11},\ldots,X_{1n_1}$, ..., $X_{r1},\ldots,X_{rn_r}$. (Z populace všech dostupných stejně starých selat téhož plemene náhodně vybereme $n_1 + n_2 + ... + n_r$ jedinců. Náhodně je rozdělíme na r souborů o rozsazích $n_1$, $n_2$, ..., $n_r$. Selata z prvního souboru podrobíme výkrmné dietě č. 1, ..., selata z r-tého souboru podrobíme výkrmné dietě č. r. Tak dostaneme realizace r nezávislých náhodných výběrů.)

Za předpokladu, že všechny náhodné výběry se řídí normálním rozložením s týmž rozptylem, můžeme testovat hypotézu o stejné účinnosti všech r výkrmných diet. (Tomuto problému je věnována 6. kapitola.)

**2.3.2. Blokové porovnávání**: Situace je charakterizována jedním náhodným výběrem $(X_{11},X_{12}, ..., X_{1r})$, ..., $(X_{n1},X_{n2}, ..., X_{nr})$ z r-rozměrného rozložení. Blokem se rozumí r-tice $(X_{i1},X_{i2}, ..., X_{ir})$, i = 1, ..., n. (Náhodně vybereme n vrhů starých selat téhož plemene a z nich vždy r sourozenců a náhodně jim přiřadíme 1. až r-tou výkrmnou dietu. Tak dostaneme realizaci náhodného výběru z r-rozměrného rozložení.)

Vyhodnocení výsledků při blokovém porovnávání se provádí např. pomocí Friedmanova testu. Jeho popis se již vymyká náplni předmětu Statistika II. Poučení lze nalézt v doporučené literatuře na str. 360.

## Kontrolní otázky

1. Popište tři způsoby plánování pokusů.
2. Jak se liší dvouvýběrové a párové porovnávání?
3. Jak se liší mnohovýběrové a blokové porovnávání?
4. Pokud u několika osob měříme krevní tlak před zátěží a po zátěži, o jaký typ porovnávání se jedná?
5. Náhodně vybereme dostatečný počet rodin s dětmi a zkoumáme, zda počet dětí ovlivňuje průměrné roční výdaje rodiny na průmyslové zboží. O jaký typ porovnávání se jedná?