

Studijní opora

POČÍTAČOVÁ STATISTICKÁ ŠETŘENÍ

Viktor Hynčica, prom.mat.

Rozsah: 5 hodiny výuky, 10 hodin konzultací



Úvod

Předmět navazuje na přednášku Biostatistika z prvního ročníku a hlavním cílem předmětu je seznámit posluchače s používáním statistických počítačových programů ke zpracování výběrových statistických šetření na konkrétních datech. Cílem je usnadnit posluchačům zpracování závěrečných bakalářských prací, pokud jsou založeny na nějakém statistickém šetření. Úvodní přednáška vždy začíná zopakováním základních statistických pojmů :

- deskriptivní , induktivní statistika
- etapy výběrového statistického šetření
- odhady parametrů
- statistické testy hypotéz



Cílem je ukázat použití statistických programů na konkrétních datech a dát prostor i pro pomoc při zpracování studentských dat.

Obsah

1. Programy pro statistická šetření
2. Použití Excelu
3. Programy řady Epi-Info

1 . Programy pro statistická šetření

Spektrum programů pro statistické zpracování dat je široké a lze si vybrat podle náročnosti použití. Existují programové balíky pro „znalce a odborníky“ a na druhé straně i programy pro ty , kteří o statistice tolik nevědí , tedy i pro studenty. Do první skupiny patří například statistický balík programů SPSS, který v několika modulech pokrývá celou teorii statistiky.

Zkratka znamená Statistical Package for Social Science. Pro běžného uživatele je navíc i ekonomicky nepřístupný. Podobně lze hovořit i o programovém balíku BMDP – Biomedical Data Package.

Do druhé skupiny programů, určené pro běžné uživatele, lze zařadit program Excel kancelářského balíku Microsoft Office, jehož použití bohužel přináší i některá úskalí. Více lze pro běžné použití doporučit volně šiřitelné programy řady EPI-INFO, nebo program „R“. Existují i programy , které lze zařadit mezi tyto dvě skupiny, jedná se například o program STATA, Statgraphics apod..

2. Použití programu Excel

Program Excel je součástí kancelářského balíku Microsoft Office a má široké užití při jakékoli práci s daty ve formě tabulek. Program je vyvíjen řadu let a tak můžeme používat Excel ve verzi 2003, 2007, 2010 a nyní i zatím v poslední verzi 2013. Při použití pro statistická šetření je zapotřebí zvládnout použití Excelu pro založení a práci s databázemi, kdy první řádka definuje strukturu souboru a následující představují vždy jeden záznam z jednoho dotazníku. Důležité je definovat typ jednotlivých polí sloupců pro vkládané hodnoty. Pro vlastní statistické hodnocení je důležité zvládnout práci s funkcemi, kterých má Excel mnoho a jsou seskupeny do skupin podle použití, tedy i funkce pro statistiku. Nepříjemné je, že výsledek je omezen pouze na jedno políčko a tak například pro výsledek t-testu vidíme pouze významnost (**p**) a ne výslednou hodnotu t.

3. Programy řady EPI-INFO

Programy řady EPI-INFO jsou volně šiřitelné programy pro zpracování epidemiologických studií, vytvořené ve spolupráci Office Center for Disease Control v Atlantě (USA) a World Health Organization Ženeva ve Švýcarsku. Dnes jejich užití není omezeno jen na epidemiologické studie, ale jsou výborným pomocníkem pro běžná počítačová statistická zpracování v oblasti veřejného zdravotnictví ale i mimo ní. Dnes je možné používat tento program ve verzi 6 a 7. Verze 6 se stále ještě dnes používá a měla přednost v tom, že ji pracovníci Státního zdravotního ústavu přeložili do češtiny. Dnes ji není možné používat v nových operačních systémech MS Windows 7 a 8. V těchto systémech ji ale bez problému nahrazují programy EPIDATA Entry a EPIDATA Analysis, které vytvořili pracovníci „The EpiData Association“ Odense v Dánsku. Obrovská výhoda programu EPI-INFO spočívá v možnosti jak vytvořit databázi dat pro zpracování prostřednictvím jednoduchého datového formuláře, kde jednotlivé položky jsou specifikovány popisem, z kterého se generuje název položky a definovanou maskou, z které se určuje typ a rozsah položky.

Jádro klasického EPI-INFO tvoří 4 programy:

- **EPED** – textový editor pro vytváření dotazníků
- **ENTER** – zadávání či aktualizace dat
- **ANALYSIS** – statistické zpracování dat
- **CHECK** – programová kontrola zadávaných dat

Program EPIDATA Enter v sobě sdružuje programy Eped, Enter a Check. Program EPI-INFO ve verzi 7 je již plně přizpůsoben novým operačním systémům MS Windows 7 či 8 a pro ukládání dat používá databázový formát z programu Access MS Office.

Studentům je předvedeno statistické zpracování konkrétních dat programem Excel ve srovnání s programem EPIDATA

literatura

Lauritsen JM. (Ed.) EpiData Data Entry, Data Management and basic Statistical Analysis System. Odense Denmark, EpiData Association, 2000-2008.

[Http://www.epidata.dk](http://www.epidata.dk)

J.Hendl: Přehled statistických metod zpracování dat. Portál, Praha 2004.

K. Zvára: Biostatistika. Karolinum, Praha 1998.

Epi Info, verze 5 a 6, překlad anglického manuálu Andrew G. Dea,

Excel 2003, rychlémi kroky; Cronan John; 2005; ISBN:80-247-1008-0

<http://wwwn.cdc.gov/epiinfo/7/index.htm>