**OBSAH**

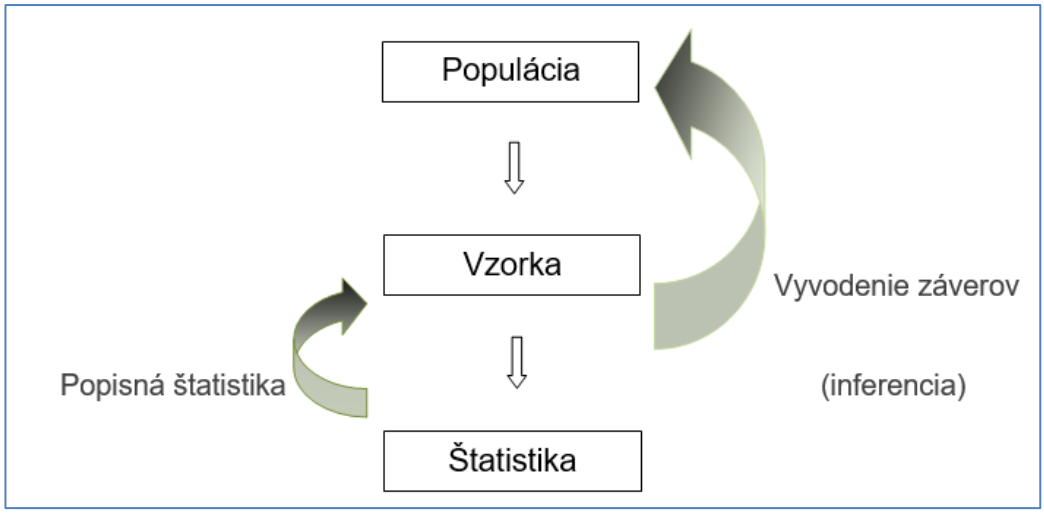
### Charakteristika a význam sumárnej štatistiky

Sumárna štatistika tvorí jeden zo základných štatistických nástrojov v rámci popisnej štatistiky pri ktorej sa najčastejšie počítajú základné štatistické charakteristiky kvantitatívnych znakov a vlastností skúmaných javov a procesov.

Sumárna štatistika by mala byť jednou s prvých analýz, ktoré použijeme ak začíname analyzovať nejaké číselné údaje. Sumárna štatistika tvorí neodmysliteľnú súčasť celej uskutočnenej štatistickej analýzy.

Štatistická analýza zahŕňa štyri základné procesy:

1. Identifikácia populácie, ktorá nás zaujíma (napr. skupina zvierat konkrétneho druhu a plemena, definovanie a výber analyzovaných znakov a vlastností).
2. Výber náhodnej vzorky a výber štatistických metód (malá pravdepodobnoť, že budeme mať k dispozícii údaje z celej populácie).
3. Štatistické výpočty na základe konkrétnych dielčich analýz slúžiace na podrobný popis a analýzu náhodnej vzorky (popisná štatistika).
4. Využitie a prezentácia získaných výsledkov a informácií z náhodnej vzorky na vyvodenie všeobecných záverov (inferenčná štatistika) využiteľných na úrovni celej populácie.




Obr. 1.1 Štatistická analýza (zdroj: projekt KEGA, Candrák, 2021)

Veľmi dôležitou podmienkov výberu a použitia rôznych štatistických metód v rámci štatistických analýz je splnenie predpokladov, ktoré musia byť splnené aby konkrétna metóda bola správne použitá. Pri interpretácii výsledkov musíme preto presne vedieť, na akých predpokladoch sú jednotlivé metódy založené.

Základné štatistické charakteristiky (popisné charakteristiky) sú číselné charakteristiky, ktoré koncentrovanou formou, jedným číslom, vyjadrujú určitú vlastnosť skúmaného štatistického znaku.

Väčšinou sú použiteľné pre kvantitatívne štatistické znaky a len niektoré môžeme použiť pre kvalitatívne štatistické znaky. Predstavujú nástroj štatistiky ako prvotne pochopiť analyzované údaje.

Sumárna štatistika poskytuj v oblasti živočíšnej produkcie prvotné informácie o údajoch, ktoré môžeme použiť na zodpovedanie mnohých otázok súvisiacich s chovom, výživou, šľachtením a genetickým hodnotením zvierat. Na ich základe môžeme prijímať správne rozhodnutia o hodnotách a parametroch aktuálneho stavu hodnotených ukazovateľov a vlastností zvierat.

Popisné charakteristiky rozdeľujeme na: charakteristiky polohy, charakteristiky variability, charakteristiky šikmosti a špicatosti.

**Charakteristiky polohy:** stredné hodnoty (priemery a ostatné stredné hodnoty). Medzi priemery patria: aritmetický, geometrický, kvadratický a harmonický priemer. Priemery môžu byť vo forme jednoduchej, váženej, resp poznáme aj priemery kĺzavé. Medzi ostatné stredné hodnoty patria: modus a medián.

**Charakteristiky variability:** miery variability, ktorých veľkosť ovplyvňujú len niektoré hodnoty znaku v súbore a miery variability, ktorých veľkosť ovplyvňujú každá hodnota znaku v súbore.

Do prvej skupiny patria: variačné rozpätie, kvantilové rozpätie, kvartilové rozpätie, kvartilová odchýlka. Do druhej skupiny patria: absolútne charakteristiky (priemerná odchýlka, rozptyl, smerodajná-štandardná odchýlka) a relatívne charakteristiky (pomerná priemerná odchýlka a variančný koeficient).

**Charakteristiky šikmosti a špicatosti:** koeficienty, ktoré popisujú šikmosť, alebo špicatosť spravidla normálneho rozdelenia početnosti znaku. Medzi tieto charakteristiky patria: koeficient šikmosti a koeficient špicatosti.

**Praktický postup základného popisu a skúmania údajov:**

1. Výber vhodnej metódy (metód) popisu údajov (zohľadnenie typu analyzovaných údajov).
2. Prvotné grafické zobrazenia údajov (výber vhodného grafického zobrazenia).
3. Výpočet vybratých štatistických charakteristík (sumárna štatistika).
4. Interpretácia a prezentácia výsledkov sumárnej štatistiky.
5. Základná distribučná analýza (analýza rozdelenia početnosti).
6. Interpretácia a prezentácia základnej distribučnej analýzy (normálne rozdelenie).

1. Výber vhodnej metódy (metód) popisu údajov (zohľadnenie typu analyzovaných údajov).

Rozlišujeme dva základné typy údajov: údaje spojité a údaje diskrétne. Typ údajov určuje jednoznačne spôsob ako údaje popisovať a sumarizovať ich výsledky (iný spôsob je vprípade údajov spojitých a iný v prípade údajov diskrétnych).

Spojité údaje (vek, hmotnosť, produkcia mlieka): premenné majú neobmedzený počet možných hodnôt v rámci daného rozsahu, hodnoty sú číselné. Spojité údaje sa tiež nazývajú aj intervalové údaje.

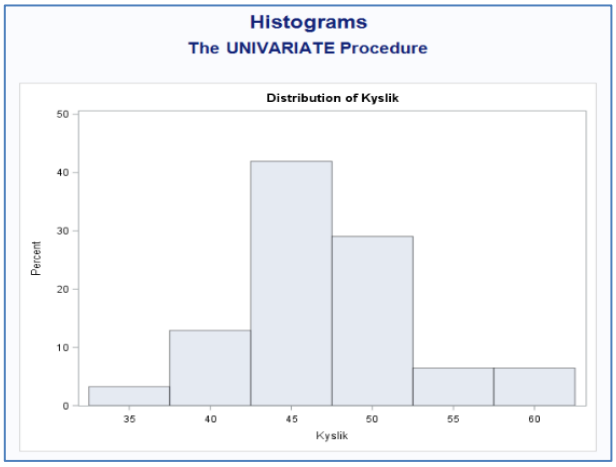
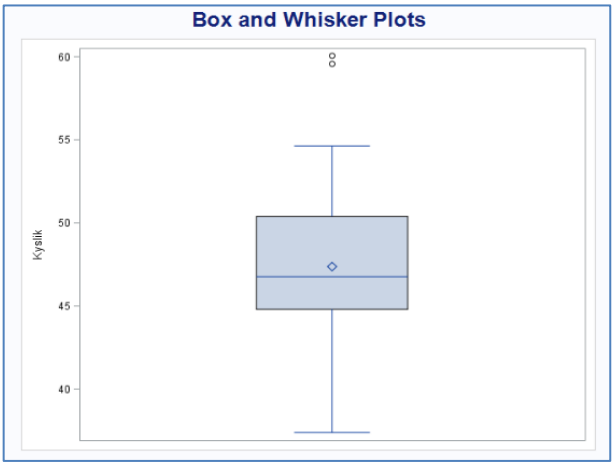
Diskrétne údaje (pohlavie farba, genotype, bodová klasifikácia, Áno/Nie) majú zvyčajne malý počet odlišných hodnôt v rámci daného rozsahu. Hodnoty môžu byť znakové alebo číselné. Diskrétne údaje sa tiež nazývajú aj kategorické (kategoriálne) údaje.

2. Prvotné grafické zobrazenia údajov (výber vhodného grafického zobrazenia).

Začiatkom každej analýzy údajov je ich kontrola z hľadiska toho, či neobsahujú nejaké chyby, nelogické resp. nepravdivé hodnoty a stavy. Najjednoduchší spôsob, ako skontrolovať, alebo preskúmať svoje údaje z tohto pohľadu, je jednoduché grafické zobrazenie ich hodnôt.

Z hľadiska tohto účelu ako základné grafické zobrazenia sa používajú: histogramy, krabicové grafy a stĺpcové grafy.

Tab 1.1 Grafické zobrazenia (zdroj: projekt KEGA, Candrák, 2021)

| | |
|---|---|
|  |  |
| Histogram | Krabicový graf |

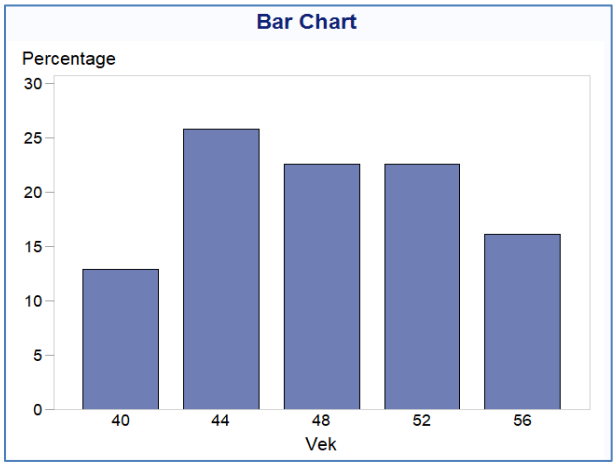
Histogram zobrazuje rozdelenie (distribúciu) spojitej premennej. Distribúcia je súbor dátových hodnôt usporiadaných v poradí spolu s relatívnou frekvenciou. Histogram poskytuje všeobecný grafický obraz o hodnotenej premennej. Môžeme vidieť, či sú údaje vycentrované alebo viac rozložené na jednej strane. V tomto príklade máme niektoré hodnoty, ktoré sú mierne vychýlené

a preto bude potrebné vykonať ich detailnejšie posúdenie pomocou samostatnej distribučnej analýzy.

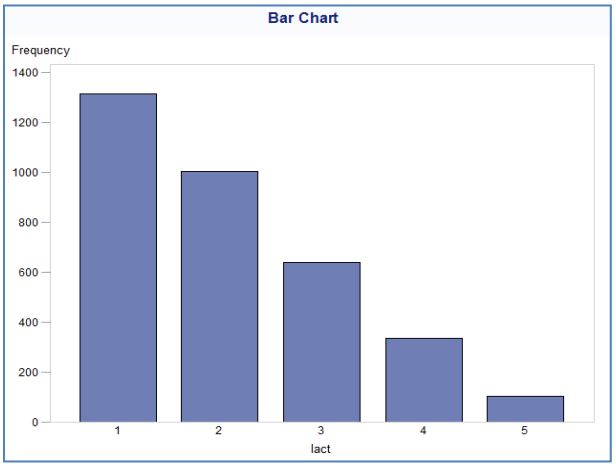
Krabicové grafy poskytujú veľmi užitočné informácie o variabilite údajov a extrémnych hodnotách niektorých údajov. Graf je založený na kvartiloch, alebo percentiloch analyzovaných údajov. Kvartily (percentily) zobrazujú a predstavujú pozíciu v údajoch, ktorá je väčšia ako daný podiel hodnôt údajov. Bežne uvádzané percentilové hodnoty sú: 25. percentil (spodný okraj obdĺžnika), nazývaný prvý kvartil, 50. percentil, nazývaný medián (stredová čiara) a 75. percentil, nazývaný tretí kvartil (horný okraj obdĺžnika). Stredová značka predstavuje aritmetický priemer. Body mimo grafu predstavujú extrémne hodnoty.

Najčastejšie sú v stĺpcových grafoch zobrazované diskrétne, alebo kategorizované premenné. V tomto príklade (Tabuľka 1.2) máme zobrazené početnosti osôb v percentách podľa stanovených vekových skupín (graf 1) a početnosti dojníc pdľa laktácií (graf 2).

Tab 2.2 Grafické zobrazenia (zdroj: projekt KEGA, Candrák, 2021)




Stĺpcový graf 1




Stĺpcový graf 2

3. Výpočet vybratých štatistických charakteristík (sumárna štatistika).

Výpočet základných resp. vybratých štatistických charakteristík predstavuje v skutočnosti výpočet mier centrálnej tendencie údajov, ktoré sa týkajú lokalizácie stredu distribúcie hodnôt a výpočet mier variability.

4. Interpretácia a prezentácia výsledkov sumárnej štatistiky.

Najčastejšie používanými mieramy centrálnej tendencies sú: aritmetický priemer, modus a median. Aritmetický priemer možno vypočítať len na spojitých údajoch.

Pre medián musíme zoradiť hodnoty od najnižšej po najvyššiu a vybrať strednú hodnotu. Ak existuje párny počet pozorovaní, potom použijeme priemer dvoch stredných hodnôt. Medián môžeme vypočítať na spojitých údajoch a tiež aj na diskrétnych údajoch, ak existuje nejaké prirodzené zoradenie ich hodnôt.

Modus je najbežnejšou (najčastejšie sa opakujúcou) hodnotou premennej vo výberovom súbore. Modus je možné získať pre spojité aj pre diskrétne údaje. Modus nemusí mať vždy riešenie (žiadna konkrétna hodnota sa nevyskytuje najviac).

Najpoužívanejšími mieramy variability sú: variačné rozpätie, kvartilové rozpätie (3. kvartil mínus 1. kvartil), rozptyl, smerodajná odchyľka a variačný koeficient. Variačné rozpätie je najjednoduchšie vypočítať a je to najväčšia hodnota mínus najmenšia. Táto miera variability ale závisí iba od dvoch extrémnych hodnôt a v praxi nemá až takú vypovedaciu silu. Kvartilové rozpätie je v skutočnosti rozpätie stredných 50 % údajov, a je menej citlivé na extrémne hodnoty údajov.

Rozptyl je v skutočnosti priemerný štvorec rozdielu medzi každou hodnotou premennej a jej priemerom. Je to najvýznamnejšia miera variability, ale pretože to je mocninový výraz, rozptyl je ťažké interpretovať. Smerodajná odchýlka je druhá odmocnina rozptylu, čo znamená, že je pre nás oveľa jednoduchšie interpretovateľná a preto je najpoužívanejšou mierou variability.

Smerodajná odchýlka závisí od mierky, na ktorej sa merajú údaje, takže štandardná odchýlka hmotnosti nameraná v kilogramoch sa bude líšiť od štandardnej odchýlky hmotnosti nameranej v gramoch. Smerodajná odchýlka sa nedá preto použiť pri porovnávaní variability znakov roznych mierok. Ak chceme porovnať variabilitu premenných v rôznych mierkach merania, používame variačný koeficient.

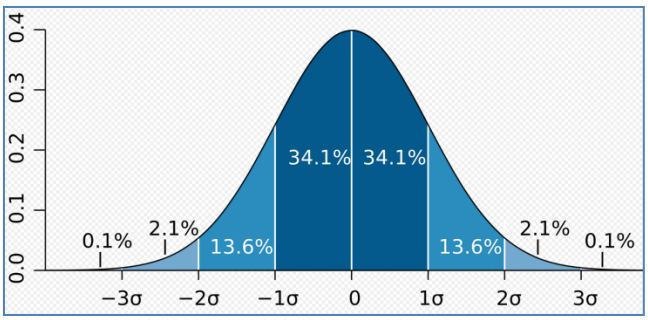
5. Základná distribučná analýza (analýza rozdelenia početnosti).

Tvar distribúcie (tvar distribučnej krivky) je dôležitý na určenie, či miery centrálnej tendencie a miery variability, ktoré sme použili sú vhodne zvolené. Hodnotíme skreslenosť, alebo neskreslenosť rozdelenia početnosti analyzovaných údajov. Jedná sa o posúdenie toho, či rozdelenie početnosti je symetrické, alebo asymetrické. V oblasti biologického výskumu najčastejšie posudzujeme, či rozdelenie početnosti má normálneho rozdelenia (Gausova krivka normálneho rozdelenia).

6. Interpretácia a prezentácia základnej distribučnej analýzy (normálne rozdelenie).

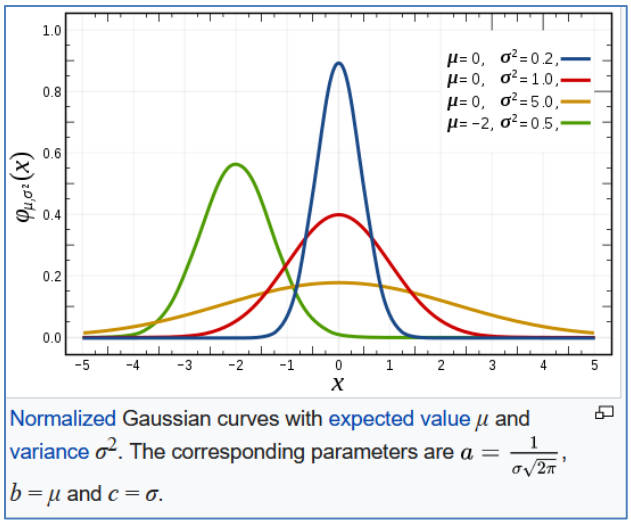
Normálne rozdelenie početnosti je najviac používaným rozdelením početnosti využívaným v štatistických analýzach v rôznych oblastiach praktickej a výskumnej činnosti biologického aj nebiologického charakteru. Normálne rozdelenie je kompletne popísané priemerom a smerodajnou

odchýlkou. Typická krivka normálneho rozdelenia má súmerný zvoncový tvar. Jedným z dôležitých predpokladov väčšiny používaných štatistických metód je predpoklad zachovania normality údajov.




Obr. 1.2 Gaussova krivka normálneho rozdelenia

zdroj: https://cs.wikipedia.org/wiki/Normální_rozdělení

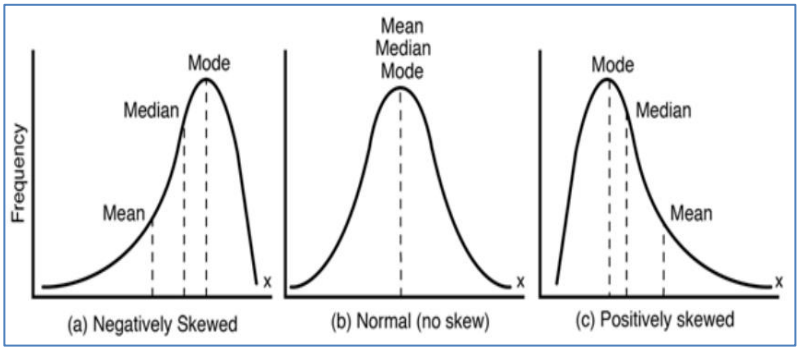



Obr. 1.3 Normalizované Gaussove krivky normálneho rozdelenia

zdroj: https://en.wikipedia.org/wiki/Gaussian_function

Ak analyzované údaje majú normálne rozdelenie, potom približne platí, že 68 % údajov patrí do rozpätia jednej smerodajnej odchýlky od priemeru, 95 % údajov patrí do rozpätia dvoch smerodajných odchýliek od priemeru a 99 % údajov patrí do rozpätia troch smerodajných odchýliel od priemeru. Uvedené pravdepodobnosti nám umožňujú urobiť závery o populácii z vybratej vzorky údajov.

Normálne rozdelenie je symetrické podľa rovnakého priemeru a mediánu. Iné rozdelenia ale môžu vykazovať rôzne tvary, ktoré možno merať šikmosťou.




Obr. 1.4 Tvary rozdelenia početnosti Obr. 1.4 Tvary rozdelenia početnosti
(zdroj: Durkhure and Lodwal, 2014)

Pozitívna symetria (ľavostranná) nastáva vtedy, keď je priemer je menší ako medián, naopak negatívna symetria (pravostranná) nastane ak je priemer väčší ako medián.

**Vzorové údaje 1.1** (zdroj: Databáza údajov projektu KEGA)

Vzorové údaje 1.1 pre výpočet a interpretáciu výsledkov sumárnej štatistiky tvoria základné ukazovatele mliekovej úžitkosti vybratej skupiny kráv slovenského strakatého plemena v rámci jedného poľnohospodárskeho podniku v Slovenskej republike. Ukazovatele mliekovej úžitkovosti sú zaznamenané na úrovni normovaných laktácií, Sledované boli prvé tri normované laktácie.

Súbor údajov obsahuje nasledovné premenné:

| | |
|---|---|
| **CISLO** | číslo zvieraťa |
| **PL** | poradie laktácie |
| **OTEC** | línia-register otca zvieraťa |
| **PLEM** | plemenný typ |
| **KODV** | kód vyradenia zvieraťa |
| **ROK** | rok otelenia |
| **LDNI** | laktačné dni |
| **MLIEKO** | produkcia mlieka (kg) |
| **TUK** | produkcia tuku (kg) |
| **TUKP** | obsah tuku (%) |
| **BIELK** | produkcia bielkovín (kg) |
| **BIELKP** | obsah bielkovín (%) |
| **LAKT** | produkcia laktózy (kg) |
| **LAKTP** | obsah laktózy (%) |

Vzor súboru údajov (10 záznamov)

| CISLO | PL | OTEC | PLEM | KODV | ROK | LDNI | MLIEKO | TUK | TUKP | BIELK | BIELKP | LAKT | LAKTP |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 000003950 | 02 | KF081 | S0 | 54 | 2000 | 305 | 4393 | 152 | 3.46 | 140 | 3.19 | 210 | 4.78 |
| 000005026 | 02 | TB001 | S2 | 00 | 2003 | 305 | 4000 | 119 | 2.98 | 126 | 3.15 | 194 | 4.85 |
| 000005856 | 02 | HX032 | S0 | 54 | 2000 | 305 | 4972 | 239 | 4.81 | 160 | 3.22 | 222 | 4.47 |
| 000005866 | 03 | DSO001 | S0 | 00 | 2003 | 305 | 4454 | 232 | 5.21 | 154 | 3.46 | 210 | 4.71 |
| 000008866 | 02 | RBB001 | S0 | 00 | 2003 | 305 | 4793 | 199 | 4.15 | 154 | 3.21 | 226 | 4.72 |
| 000009950 | 03 | KF081 | SC | 54 | 2003 | 305 | 4147 | 161 | 3.88 | 138 | 3.33 | 194 | 4.68 |
| 000013950 | 02 | NEZ000 | S2 | 54 | 2000 | 269 | 4368 | 137 | 3.14 | 139 | 3.18 | 213 | 4.88 |
| 000016950 | 03 | STA002 | SC | 00 | 2001 | 305 | 4408 | 139 | 3.15 | 135 | 3.06 | 201 | 4.56 |
| 000027866 | 01 | RBB001 | S2 | 00 | 2002 | 305 | 4445 | 218 | 4.90 | 132 | 2.97 | 212 | 4.77 |

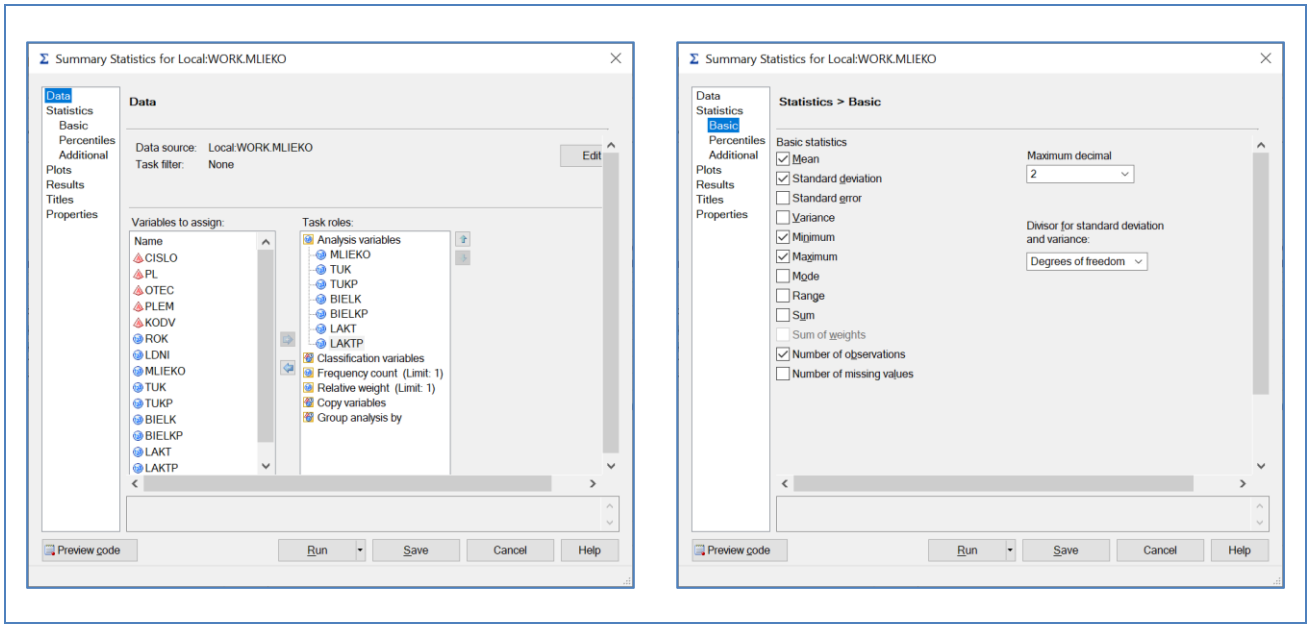
**Praktické použitie programu SAS** (SAS Enterprise Guide)

Úlohy (Tasks): Describe - Summary Statistics

**Príklad 1.1** (SAS)

Vypočítajte a popíšte základné štatistické charakteristiky ukazovateľov mliekovej úžitkovosti kráv slovenského strakatého plemena. Určite, ktorý ukazovateľ má najväčšiu a najmenšiu variabilitu v hodnotenom súbore. Pre všetky hodnotené ukazovatele zostavte a interpretujte grafy vo forme histogramov a krabicových grafov. Zistite či hodnota obsahu tuku 4,00 % a hodnota obsahu bielkovín 3,30 % sa nachádza v rozpätí konfidenčných intervalov so spoľahlivosťou 95 %.

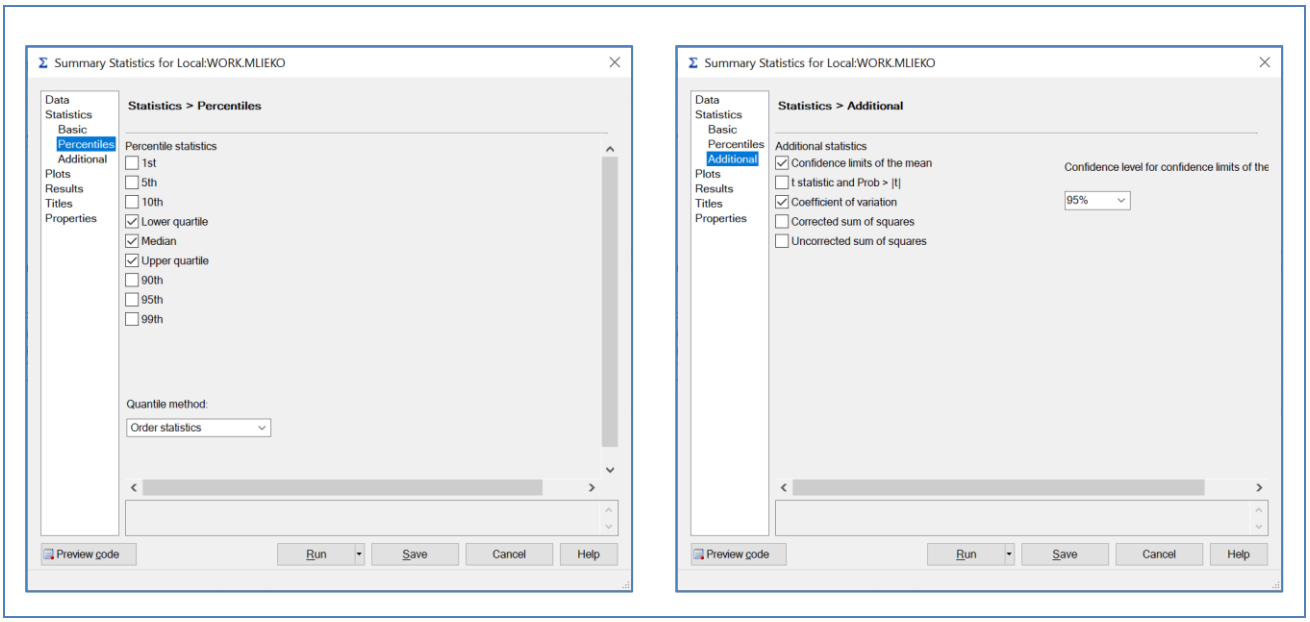
Tab 3.2a Nastavenie analýzy - príklad 1.1 (zdroj: projekt KEGA, Candrák, 2021)




Po správnom nastavení analyzovaných ukazovateľov môžeme nastaviť jednotlivé základné popisné charakteristiky, percentily a doplnkové popisné charakteristiky.

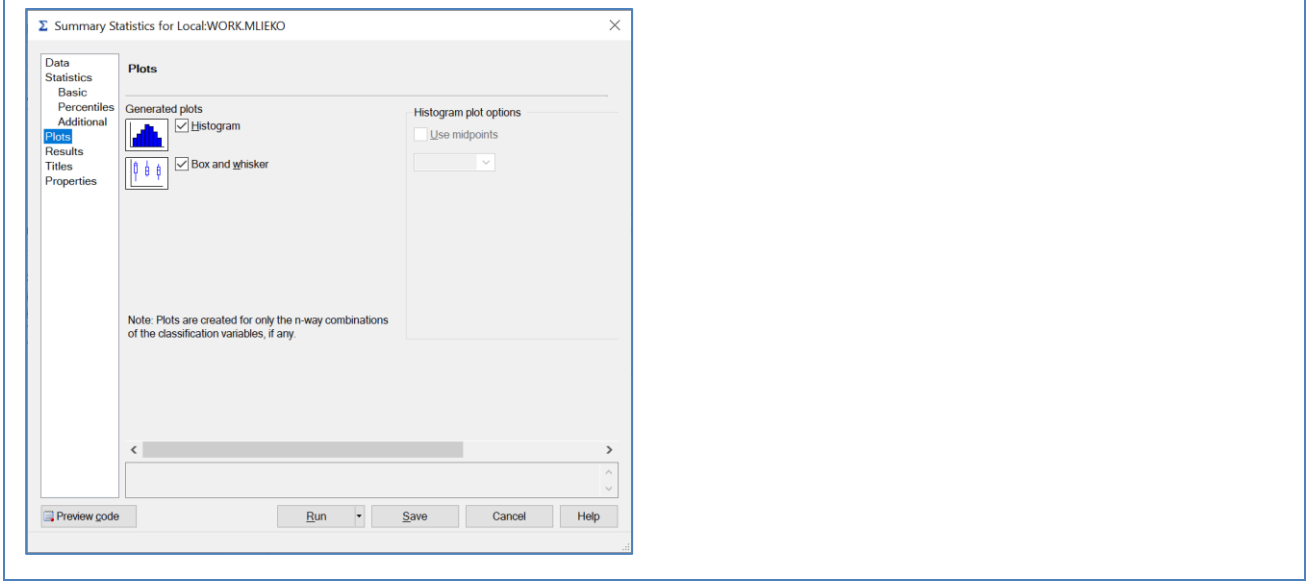
Veľmi užitočnou možnosťou je nastavenie optimálneho počtu desatinných miest pre vypočítavané stredné hodnoty, jednotlivé miery variability a vybraté doplnkové koeficienty.

Tab 4.2b Rozšírené nastavenie analýzy - príklad 1.1 (zdroj: projekt KEGA, Candrák, 2021)




Samostatne je ešte potrebné nastaviť zobrazenie požadovaných grafických výstupov (histogram a krabicový graf).

Tab 5.2c Rozšírené nastavenie analýzy - príklad 1.1 (zdroj: projekt KEGA, Candrák, 2021)



**Výsledky** (program SAS)

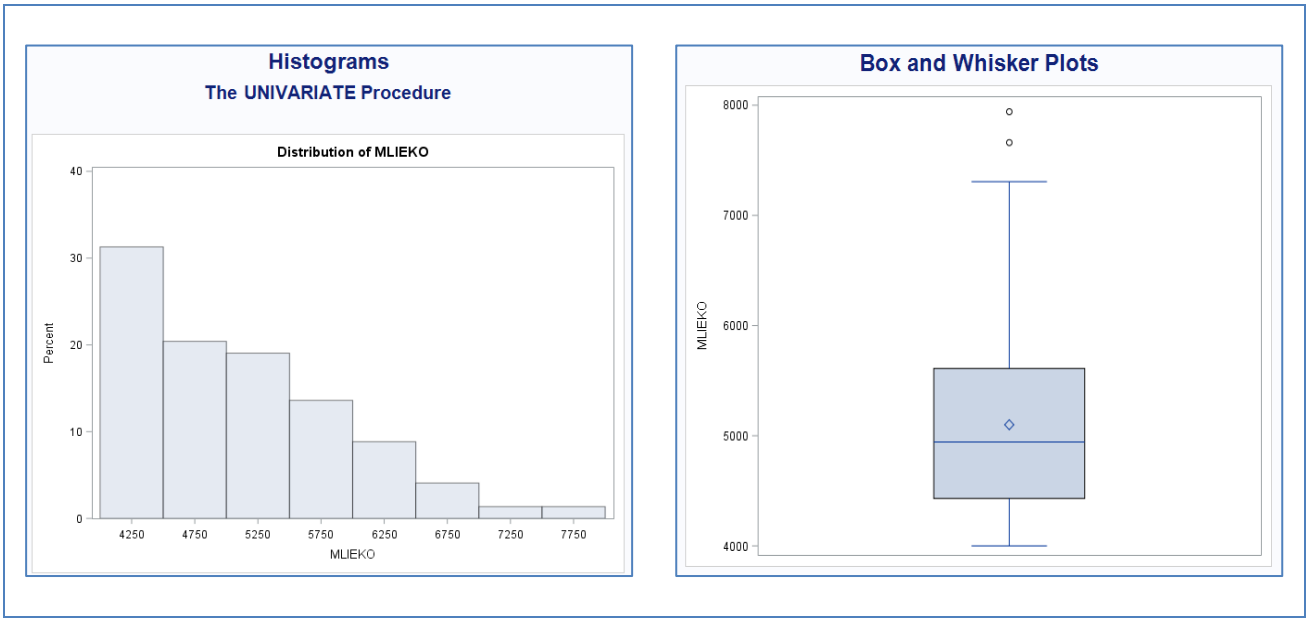
Tab 6.3 Základné štatistické charakteristiky (zdroj: projekt KEGA, Candrák, 2021)

| Variable | Mean | Std Dev | Minimum | Maximum | N | Lower Quartile | Median | Upper Quartile | Lower 95% CL for Mean | Upper 95% CL for Mean | Coeff of Variation |
|---|---|---|---|---|---|---|---|---|---|---|---|
| MLIEKO | 5099.89 | 844.00 | 4000.00 | 7941.00 | 147 | 4431.00 | 4943.00 | 5610.00 | 4962.31 | 5237.47 | 16.55 |
| TUK | 212.07 | 47.03 | 119.00 | 360.00 | 147 | 177.00 | 210.00 | 238.00 | 204.40 | 219.73 | 22.18 |
| TUKP | 4.15 | 0.60 | 2.97 | 5.56 | 147 | 3.71 | 4.12 | 4.57 | 4.06 | 4.25 | 14.39 |
| BIELK | 170.25 | 30.12 | 121.00 | 257.00 | 147 | 146.00 | 168.00 | 189.00 | 165.34 | 175.16 | 17.69 |
| BIELKP | 3.34 | 0.22 | 2.80 | 4.02 | 147 | 3.19 | 3.31 | 3.49 | 3.30 | 3.38 | 6.65 |
| LAKT | 233.61 | 44.53 | 124.00 | 371.00 | 147 | 201.00 | 224.00 | 262.00 | 226.35 | 240.87 | 19.06 |
| LAKTP | 4.58 | 0.40 | 2.21 | 5.12 | 147 | 4.54 | 4.66 | 4.79 | 4.52 | 4.65 | 8.70 |

V tabuľke 1.3 máme zobrazené všetky požadované a nastavené číselné popisné charakteristiky analyzovaného súboru mliekovej úžitkovosti kráv. Najväčšiu variabilitu (hodnota variačného koeficienta 22,12 %) sme zistili pri ukazovateli produkcia tuku v kilogramoch a naopak najnižšiu variabilitu (hodnota variačného koeficienta iba 6,65 %) sme zistili pri obsahu bielkovín.

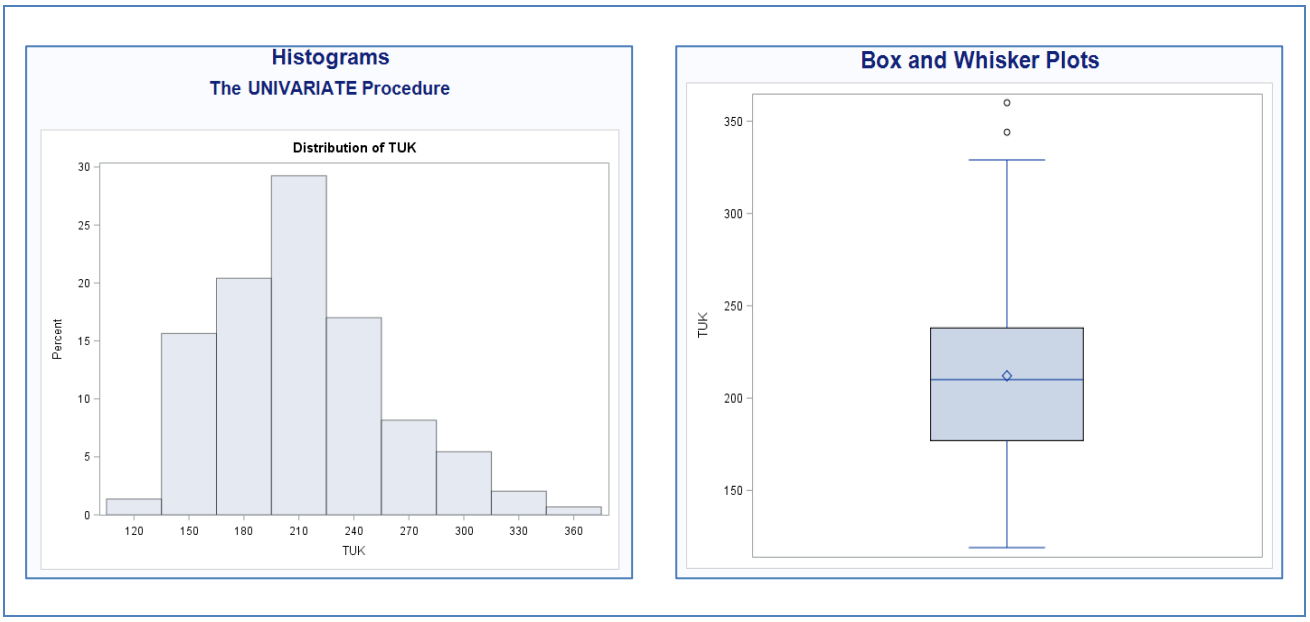
Nepotvrdil sa predpoklad, že hodnota obsahu tuku 4,00 % sa nachádza v hraniciach konfidenčného intervalu so spoľahlivosťou 95 %. Priemerná hodnota obsahu tuku (4,15 %) sa totiž nachádza vo vyšších hraniciach od 4,06 % do 4,25 %. Uvedené zistenie je ale z hľadiska mliekovej úžitkovosti kráv pozitívne. Znamená priemerné zvýšenie obsahu tuku v kravskom mlieku oproti predpokladanému obsahu tuku, čo môže mať pozitívny efekt pre konkrétneho chovateľa, aj pre samotného spracovateľa mlieka pri výrobe mliečnych výrobkov. V prípade obsahu bielkovín sa hodnota 3,30 % nachádza presne na spodnej hranice konfidenčného intervalu. Potvrdilo sa tvrdenie, že vypočítaná priemerná hodnota patrí do tohto intervalu. Grafické zobrazenia výsledkov jednotlivých ukazovateľov uvádzame v tabuľkách 1.4. - 1.10.

Tab 7.4 Mlieko (zdroj: projekt KEGA, Candrák, 2021)



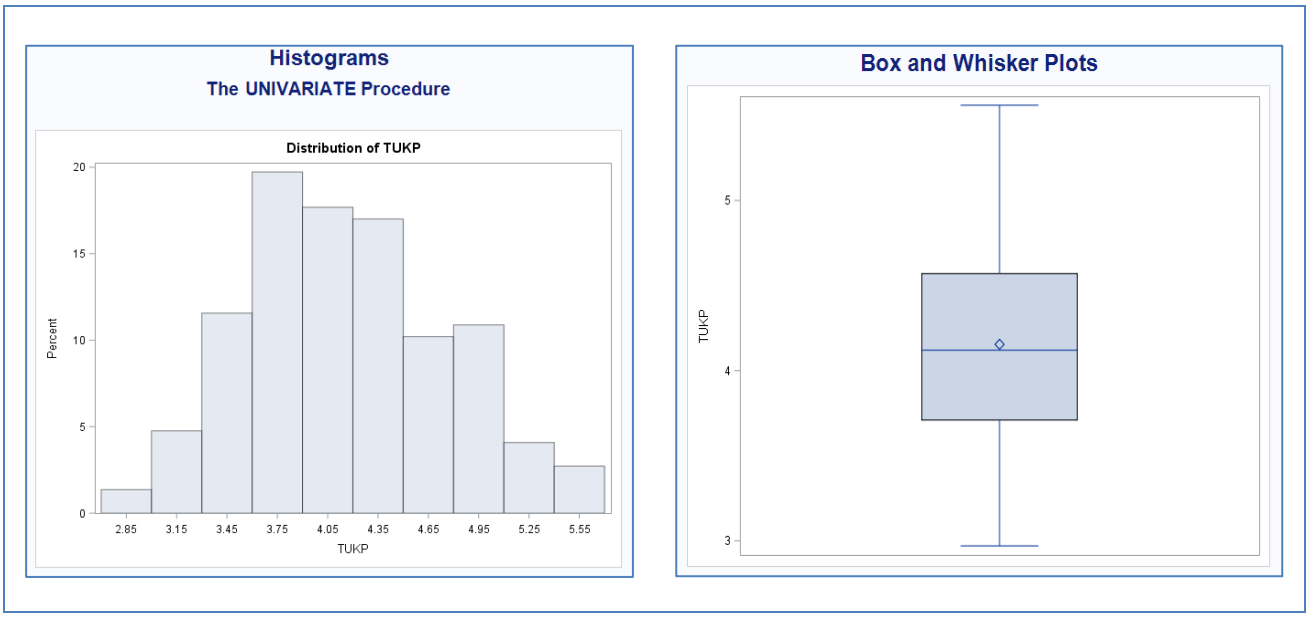
Rozdelenie početnosti ukazovateľa mlieko v kilogramoch ukazuje, že je vychýlené (zošikmené) do pravej strany (negatívna asymetria). Priemer dosahuje vyššiu hodnotu ako medián, čo je zaznamenané aj v krabicovom grafe. Podobne nie sú súmerné ani hranice, ktoré určujú vertikálne ohraničené úsečky (whiskers). Potvrdený je aj výskyt niekoľkých extrémnych hodnôt. Uvedené rozdelenie nemôžeme preto považovať za úplne normálne rozdelenie početnosti.

Tab 8.5 Tuk (zdroj: projekt KEGA, Candrák, 2021)



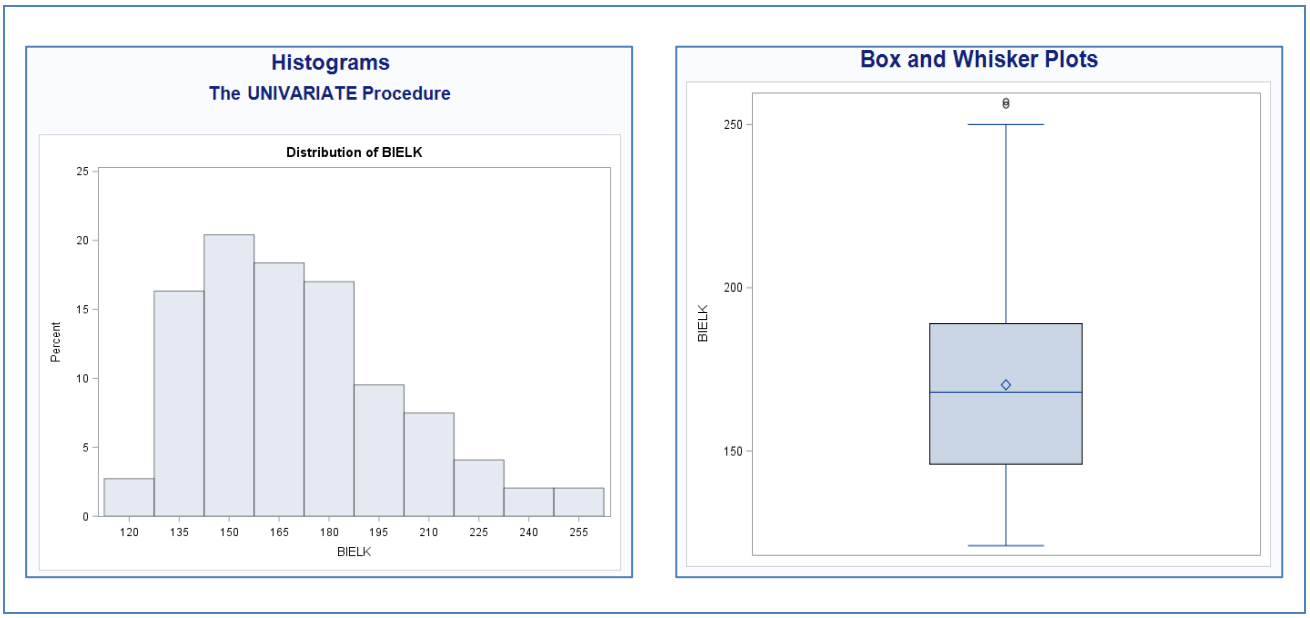

Rozdelenie početnosti ukazovateľa tuk v kilogramoch môžeme považovať za približne normálne (priemer a medián sú prakticky rovnaké). Drobná odchýlka je v spojitosti so zobrazovanými kvartilmi a niektorými extrémnymi hodnotami.

Tab 9.6 Tuk % (zdroj: projekt KEGA, Candrák, 2021)

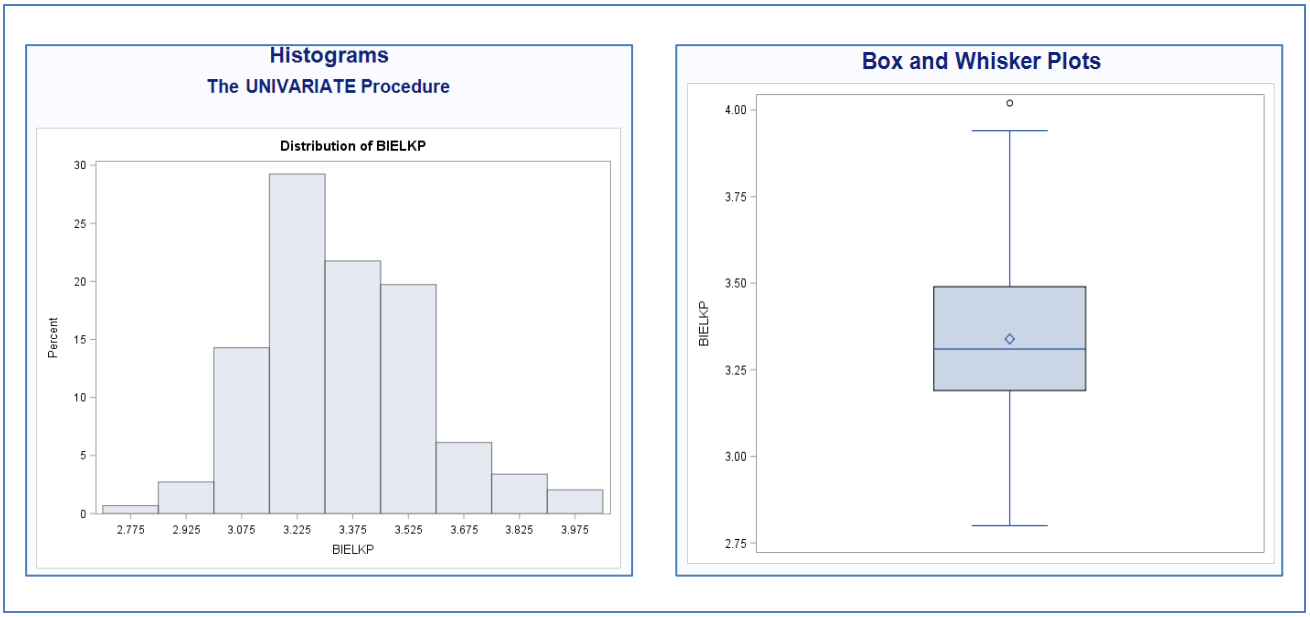


Rozdelenie početnosti ukazovateľa % tuk sa podobá normálnemu rozdeleniu. Uvedené rozdelenie môžeme preto považovať za normálne rozdelenie početnosti. Priemer a medián sú veľmi podobné. Krabicový graf to potvrdzuje tiež, neobsahuje žiadne extrémne hodnoty.

Tab 10.7 Bielkoviny (zdroj: projekt KEGA, Candrák, 2021)



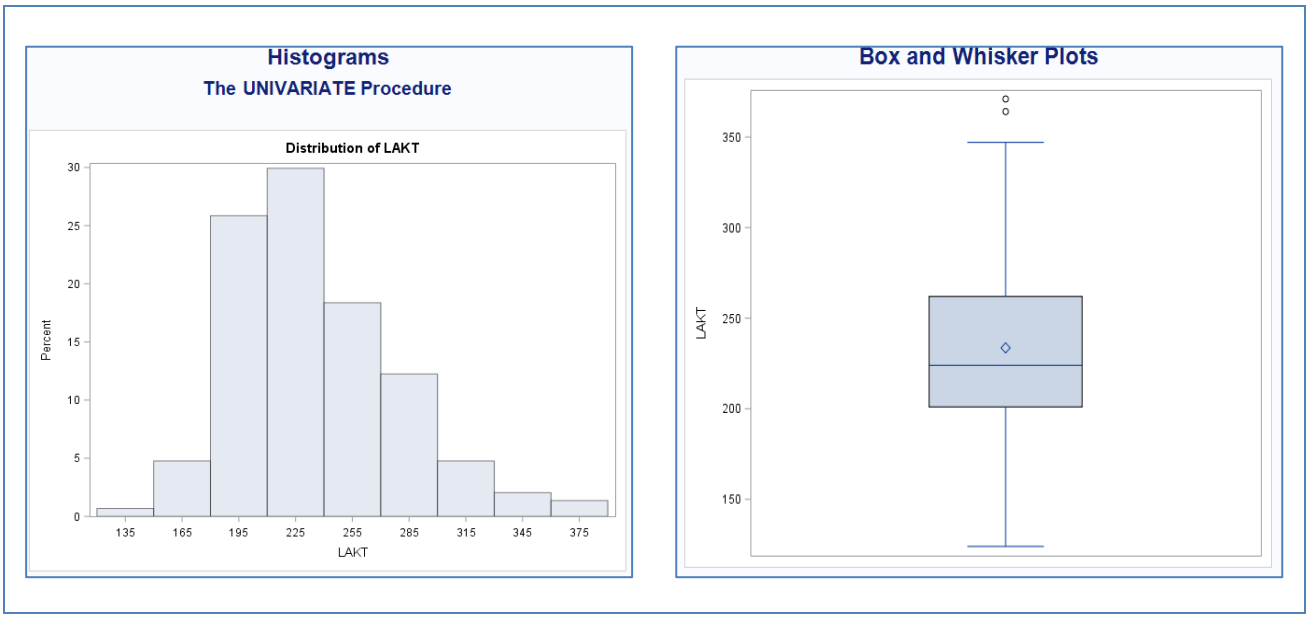

Tab 11.8 Bielkoviny % (zdroj: projekt KEGA, Candrák, 2021)




Rozdelenie početnosti ukazovateľa bielkoviny v kilogramoch môžeme tiež považovať za približne normálne (existuje ale mierna negatívna asymetria). Drobná odchýlka je v spojitosti so zobrazovanými kvartilmi a niektorými extrémnymi hodnotami.

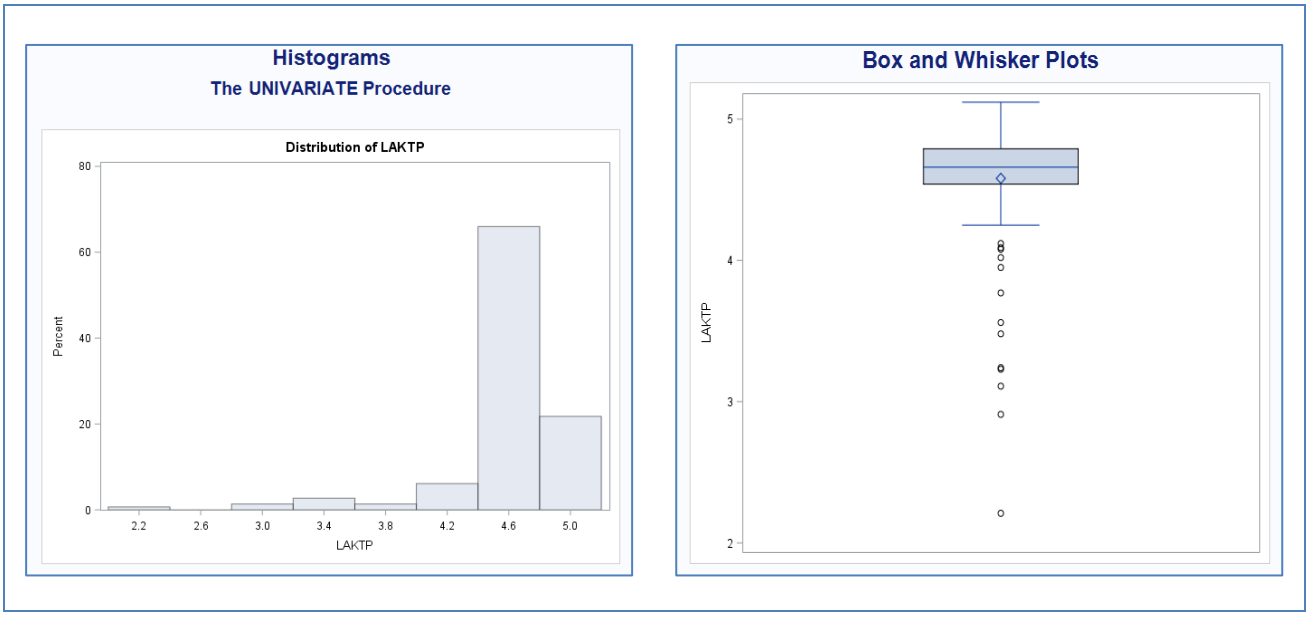
Rozdelenie početnosti ukazovateľa % bielkovín sa znovu podobá normálnemu rozdeleniu. Uvedené rozdelenie môžeme preto považovať za normálne rozdelenie početnosti. Priemer a medián sú mierne rozdielne. Krabicový graf to obsahuje jednu extrémnu hodnotu.

Tab 12.9 Laktóza (zdroj: projekt KEGA, Candrák, 2021)




Rozdelenie početnosti ukazovateľa laktóza v kilogramoch považujeme podobne ako predchádzajúce vlastnosti za približne normálne (existuje mierna negatívna asymetria). Drobná odchýlka je v rozdieloch hodnoty priemeru a mediánu a v existencii niektorých extrémnych hodnôt.

Tab 13.10 Laktóza % (zdroj: projekt KEGA, Candrák, 2021)

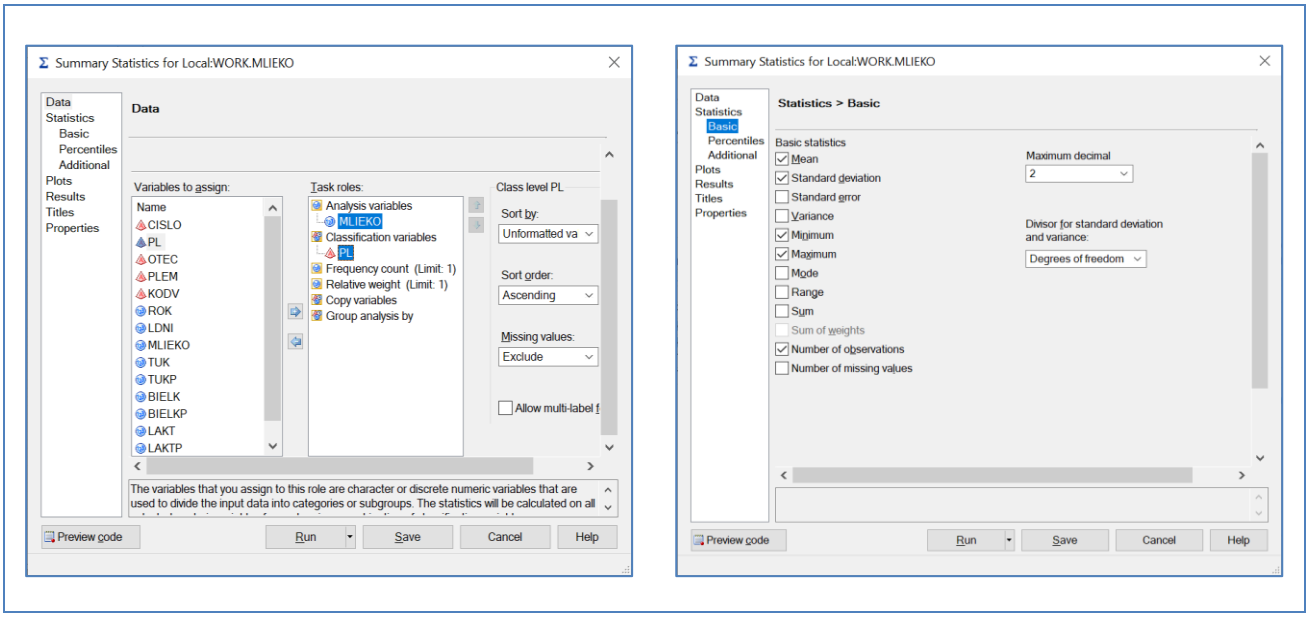


Rozdelenie početnosti ukazovateľa % laktózy ukazuje, že je vychýlené (zošikmené) do ľavej strany (pozitívna asymetria). Priemer dosahuje nižšiu hodnotu ako medián, čo je potvrdené aj v krabicovom grafe. Hoci hranice, ktoré určujú vertikálne ohraničené úsečky (whiskers) sú súmerné, existuje pomerne veľký počet minimálnych extrémnych hodnôt. Uvedené rozdelenie preto nemôžeme považovať za úplne normálne rozdelenie početnosti.

**Príklad 1.2** (SAS)

Vypočítajte a popíšte základné štatistické charakteristiky ukazovateľa produkcia mlieka v kilogramoch podľa poradia laktácie. Zistite na ktorej laktácii má produkcia mlieka najväčšiu a najmenšiu variabilitu. Pre hodnotený ukazovateľ zostavte a interpretujte grafy vo forme histogramu a vo forme krabicového grafu podľa jednotlivých laktácií. Riešenie zadanej úlohy podľa laktácií má nasledovné nastavenie (Tabuľka 1.11, ostatné nastavenia sú rovnaké ako v príklade 1.1):

Tab 14.11 Nastavenia analýzy - príklad 1.2 (zdroj: projekt KEGA, Candrák, 2021)




**Výsledky** (program SAS)

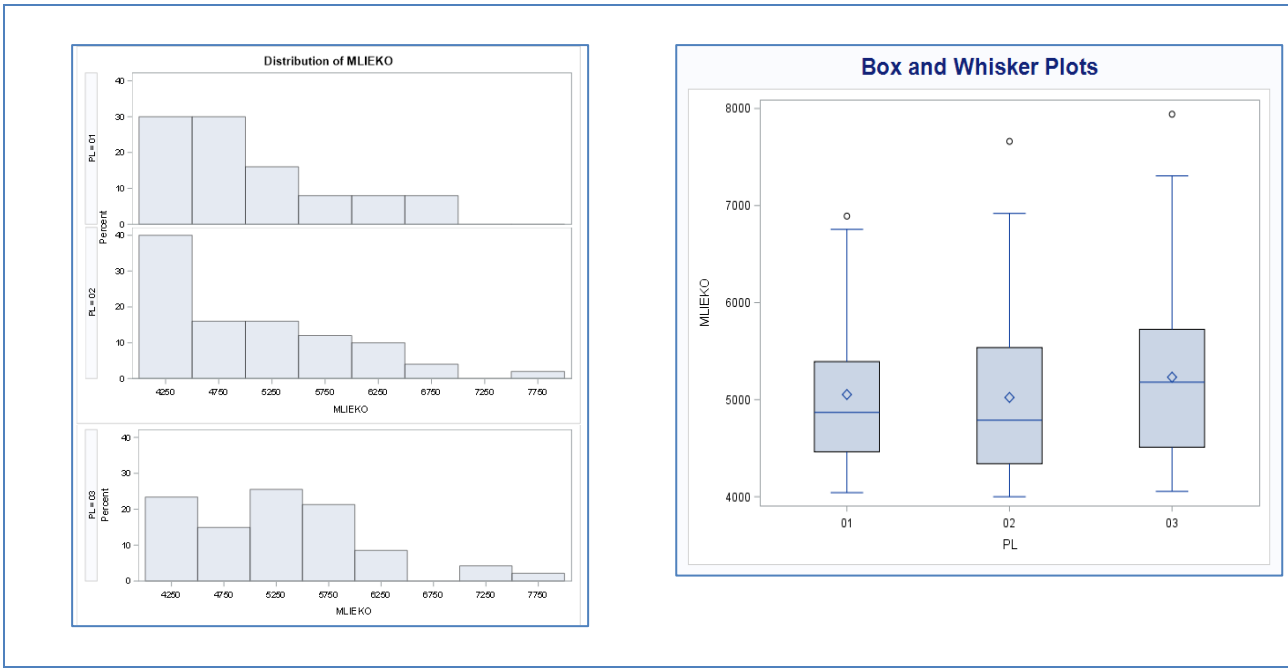
Tab 15.12 Základné štatistické charakteristiky (zdroj: projekt KEGA, Candrák, 2021)

| Analysis Variable : MLIEKO | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| PL | N Obs | Mean | Std Dev | Minimum | Maximum | N | Lower Quartile | Median | Upper Quartile | Lower 95% CL for Mean | Upper 95% CL for Mean | Coeff of Variation |
| 01 | 50 | 5052.80 | 777.48 | 4042.00 | 6891.00 | 50 | 4463.00 | 4868.50 | 5392.00 | 4831.84 | 5273.76 | 15.39 |
| 02 | 50 | 5022.98 | 897.32 | 4000.00 | 7661.00 | 50 | 4340.00 | 4788.50 | 5536.00 | 4767.97 | 5277.99 | 17.86 |
| 03 | 47 | 5231.81 | 855.98 | 4055.00 | 7941.00 | 47 | 4510.00 | 5180.00 | 5724.00 | 4980.48 | 5483.13 | 16.36 |

V tabuľke 1.12 sú uvedené základné číselné štatistické charakteristiky produkcie mlieka podľa poradia laktácie. Priemerná hodnota produkcie mlieka na prvej a druhej laktácii je približne rovnaká. Najväčšiu variabilitu dosiahla produkcia mlieka na druhej laktácii (variačný koeficient má hodnotu 17,86 %), najnižšiu variabilitu dosiahla produkcia mlieka na prvej laktácii (variačný koeficient má hodnotu 15,39 %).

Všeobecne ale môžeme tvrdiť že variabilita produkcie mlieka kráv je približne rovnaká. Grafické zobrazenia sumárnych výsledkov podľa poradia laktácie uvádza v tabuľke 1.13.

Tab 16.13 Nastavenia analýzy - príklad 1.2 (zdroj: projekt KEGA, Candrák, 2021)

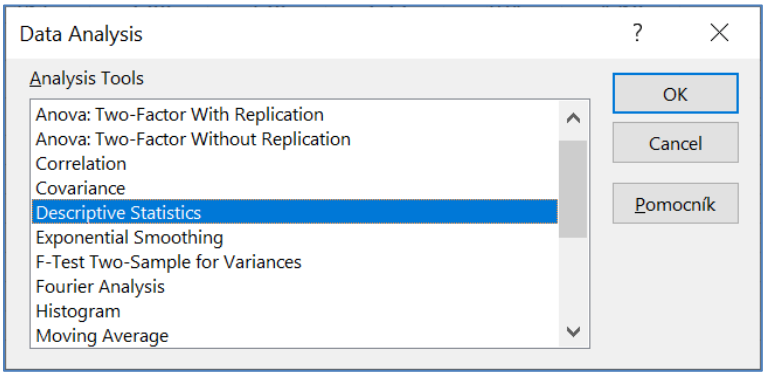



Prvé dve laktácie majú rozdelenie početnosti kilogramov mlieka veľmi podobne zošikmené do pravej strany (negatívna asymetria). Ich priemery dosahuje vyššiu hodnotu ako medián, čo je znázornené aj v krabicovom grafe. Podobne nie sú súmerné ani hranice, ktoré určujú vertikálne ohraničené úsečky (whiskers). Potvrdený je aj výskyt extrémnych hodnôt. Uvedené rozdelenia nemôžeme preto považovať za úplne normálne rozdelenia početnosti. Rozdelenie početnosti produkcie mlieka na tretej laktácie je ale rozdielne. Blíži sa skôr k normálnemu rozdeleniu početnosti. Hodnota priemeru je veľmi podobná hodnote mediánu (rozdiel je iba približne 51 kilogramov). Rozdiel priemerných hodnôt a mediánov v prípade prvých dvoch laktácií je výrazne väčší.

**Praktické použitie programu EXCEL** (Microsoft 365)

V rámci doplnkov programu Microsoft Excel je možné využiť doplnok analytických nástrojov **Data Analysis,** ktorý má jednoduchú možnosť výpočtu základných štatistických charakteristík (sumárnu štatistiku) v položke Descriptive Statistics (popisná štatistika).

Tab 17.14 Použitie analýzy údajov v programe Excel

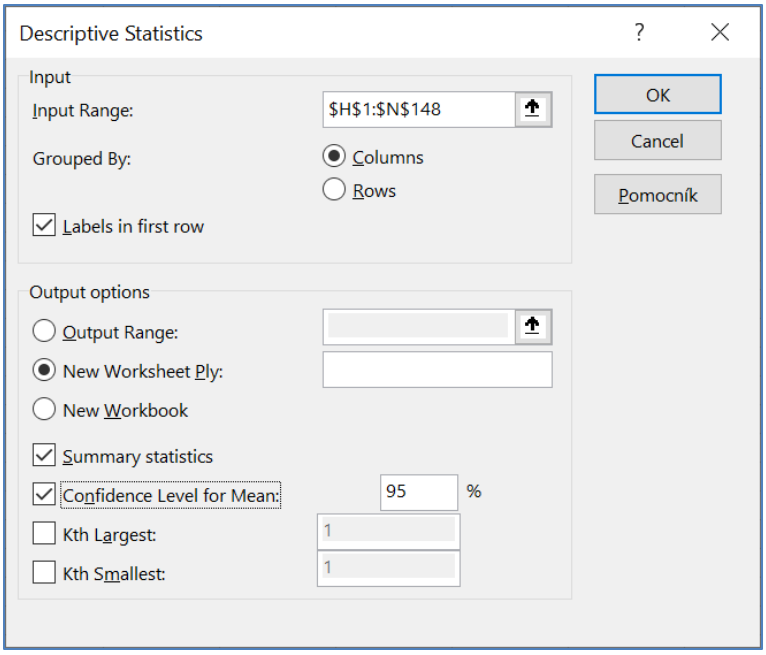



**Príklad 1.1** (Excel)

Vypočítajte a popíšte základné štatistické charakteristiky ukazovateľov mliekovej úžitkovosti kráv slovenského strakatého plemena. Určite, ktorý ukazovateľ má najväčšiu a najmenšiu variabilitu v hodnotenom súbore. Zistite či hodnota obsahu tuku 4,00 % a hodnota obsahu bielkovín 3,30 % sa nachádza v rozpätí konfidenčných intervalov so spoľahlivosťou 95 %.

Detailné nastavenie vstupných (analyzovaných) údajov a odporúčané nastavenie ostatných parametrov výpočtu (tabuľka 1.15):

Tab 18.15 Detainé nastavenie analýzy údajov v programe Excel



Uvedené nastavenie predpokladá, že všetky hodnotené ukazovatele sú umiestnené v stĺpcoch vedľa seba a majú v prvom riadku svoje označenie (popisky).

**Výsledky** (Excel)

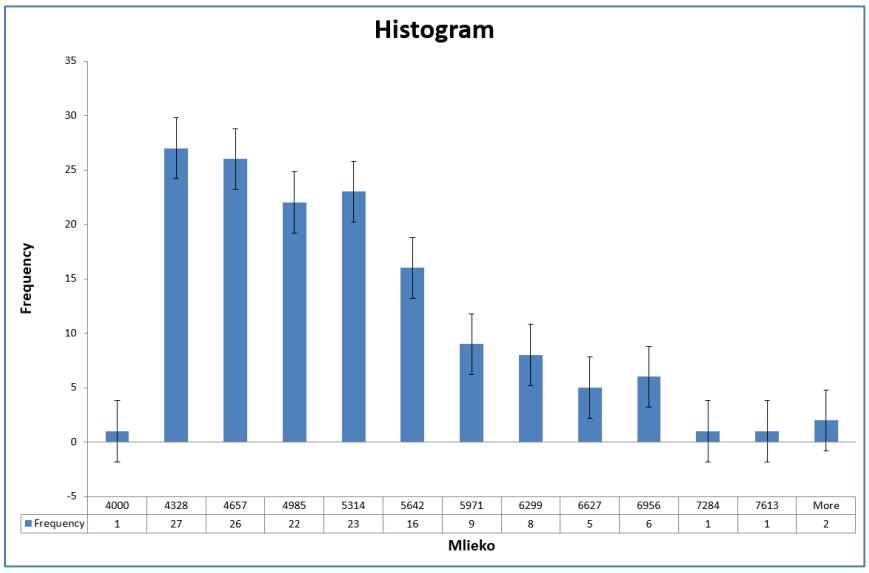
V tabuľke 16 uvádzame základné štatistické charakteristiky hodnotených ukazovateľov mliekovej úžitkovosti.

Tab 19.16 Základné štatistické charakteristiky (zdroj: projekt KEGA, Candrák, 2021)

| *MLIEKO* | | *TUK* | | *TUKP* | | *BIELK* | | *BIELKP* | |
|---|---|---|---|---|---|---|---|---|---|
| Mean | 5099.891 | Mean | 212.068 | Mean | 4.154422 | Mean | 170.2517 | Mean | 3.338844 |
| Standard Error | 69.61228 | Standard Error | 3.879146 | Standard Error | 0.049297 | Standard Error | 2.484222 | Standard Error | 0.018319 |
| Median | 4943 | Median | 210 | Median | 4.12 | Median | 168 | Median | 3.31 |
| Mode | 5129 | Mode | 215 | Mode | 4.28 | Mode | 180 | Mode | 3.43 |
| Standard Deviation | 844.0041 | Standard Deviation | 47.03215 | Standard Deviation | 0.597691 | Standard Deviation | 30.11959 | Standard Deviation | 0.222105 |
| Sample Variance | 712342.9 | Sample Variance | 2212.023 | Sample Variance | 0.357234 | Sample Variance | 907.1896 | Sample Variance | 0.049331 |
| Kurtosis | 0.500828 | Kurtosis | 0.338424 | Kurtosis | -0.45604 | Kurtosis | 0.229075 | Kurtosis | 0.199029 |
| Skewness | 0.939199 | Skewness | 0.626027 | Skewness | 0.350831 | Skewness | 0.783495 | Skewness | 0.405111 |
| Range | 3941 | Range | 241 | Range | 2.59 | Range | 136 | Range | 1.22 |
| Minimum | 4000 | Minimum | 119 | Minimum | 2.97 | Minimum | 121 | Minimum | 2.8 |
| Maximum | 7941 | Maximum | 360 | Maximum | 5.56 | Maximum | 257 | Maximum | 4.02 |
| Sum | 749684 | Sum | 31174 | Sum | 610.7 | Sum | 25027 | Sum | 490.81 |
| Count | 147 | Count | 147 | Count | 147 | Count | 147 | Count | 147 |
| Confidence Level(95.0%) | 137.5779 | Confidence Level(95.0%) | 7.666533 | Confidence Level(95.0%) | 0.097427 | Confidence Level(95.0%) | 4.909681 | Confidence Level(95.0%) | 0.036205 |

Výsledky sú identické ako v prípade použitia programu SAS (SAS Enterprise Quide). Rozdiel je iba v odlišnom vyjadrení konfidenčného intervalu, ktorý je uvedený vo forme polovice jeho rozsahu intervalu a nie ako spodná a horná hranica intervalu. Uvedenú hodnotu musíme odpočítať a pripočítať k aritmetickému priemeru (+- Mean).

Grafické zobrazenie rozdelenia početnosti ukazovateľa mlieko v kilogramoch, tak ako bolo vytvorené v prostredí programu Excel uvádzame v obrázku 1.5.




Obr. 1.5 Rozdelenia početnosti - mlieko

Grafické zobrazenie bolo tiež vytvorené v rámci analýzy údajov: Data Analysis - Histogram. V grafe sú okrem rozdelenia početnosti zobrazené aj chybové úsečky (štandardná chyba) v rámci jednotlivých skúpín rozdelenia početnosti. Program Excel umožňuje v rámci zostavovania grafov použiť a zostaviť aj iné formy zobrazenia výsledkov popisnej štatistiky a rozdelenia početnosti údajov.

**Príklad 1.2** (Excel)

Vypočítajte a popíšte základné štatistické charakteristiky ukazovateľa produkcia mlieka v kilogramoch podľa poradia laktácie. Zistite na ktorej laktácii má produkcia mlieka najväčšiu a najmenšiu variabilitu.

Úloha predpokladá usporiadanie všetkých záznamov podľa poradia laktácie následne vykonanie troch samostatných analýz popisnej štatistiky v module Data Analysis.

**Výsledky** (Excel)

V tabuľkách 1.17 - 1.19 uvádzame vypočítane základné štatistické charakteristiky hodnotených ukazovateľov mliekovej úžitkovosti kráv podľa laktácií

Tab 20.17 Základné štatistické charakteristiky - 1. laktácia (zdroj: projekt KEGA, Candrák, 2021)

| *MLIEKO* | | *TUK* | | *TUKP* | | *BIELK* | | *BIELKP* | |
|---|---|---|---|---|---|---|---|---|---|
| Mean | 5052.8 | Mean | 209.66 | Mean | 4.1604 | Mean | 171.18 | Mean | 3.3894 |
| Standard Error | 109.953 | Standard Error | 5.52535 | Standard Error | 0.08421 | Standard Error | 3.92363 | Standard Error | 0.03011 |
| Median | 4868.5 | Median | 215 | Median | 4.225 | Median | 171 | Median | 3.395 |
| Mode | 4898 | Mode | 215 | Mode | 4.28 | Mode | 147 | Mode | 3.43 |
| Standard Deviation | 777.483 | Standard Deviation | 39.0701 | Standard Deviation | 0.59545 | Standard Deviation | 27.7442 | Standard Deviation | 0.21289 |
| Sample Variance | 604480 | Sample Variance | 1526.47 | Sample Variance | 0.35456 | Sample Variance | 769.742 | Sample Variance | 0.04532 |
| Kurtosis | 0.00768 | Kurtosis | -0.3333 | Kurtosis | -0.5591 | Kurtosis | 0.08851 | Kurtosis | 0.17989 |
| Skewness | 0.96576 | Skewness | -0.1593 | Skewness | 0.24463 | Skewness | 0.66987 | Skewness | 0.5288 |
| Range | 2849 | Range | 176 | Range | 2.57 | Range | 118 | Range | 0.97 |
| Minimum | 4042 | Minimum | 120 | Minimum | 2.97 | Minimum | 132 | Minimum | 2.97 |
| Maximum | 6891 | Maximum | 296 | Maximum | 5.54 | Maximum | 250 | Maximum | 3.94 |
| Sum | 252640 | Sum | 10483 | Sum | 208.02 | Sum | 8559 | Sum | 169.47 |
| Count | 50 | Count | 50 | Count | 50 | Count | 50 | Count | 50 |
| Confidence Level(95. | 220.958 | Confidence Level(95. | 11.1036 | Confidence Level(95. | 0.16923 | Confidence Level(95. | 7.88482 | Confidence Level(95 | 0.0605 |

Tab 21.18 Základné štatistické charakteristiky - 2. laktácia (zdroj: projekt KEGA, Candrák, 2021)

| *MLIEKO* | | *TUK* | | *TUKP* | | *BIELK* | | *BIELKP* | |
|---|---|---|---|---|---|---|---|---|---|
| Mean | 5022.98 | Mean | 212 | Mean | 4.1958 | Mean | 165.28 | Mean | 3.2926 |
| Standard Error | 126.9 | Standard Error | 8.12439 | Standard Error | 0.09459 | Standard Error | 4.35356 | Standard Error | 0.02992 |
| Median | 4788.5 | Median | 195.5 | Median | 4.065 | Median | 153.5 | Median | 3.265 |
| Mode | #N/A | Mode | 239 | Mode | 5.56 | Mode | 132 | Mode | 3.18 |
| Standard Deviation | 897.318 | Standard Deviation | 57.4481 | Standard Deviation | 0.66884 | Standard Deviation | 30.7843 | Standard Deviation | 0.21154 |
| Sample Variance | 805179 | Sample Variance | 3300.29 | Sample Variance | 0.44734 | Sample Variance | 947.675 | Sample Variance | 0.04475 |
| Kurtosis | 0.31039 | Kurtosis | 0.19901 | Kurtosis | -0.5669 | Kurtosis | 0.00252 | Kurtosis | 1.94619 |
| Skewness | 0.96523 | Skewness | 0.97603 | Skewness | 0.42622 | Skewness | 0.89469 | Skewness | 0.96321 |
| Range | 3661 | Range | 241 | Range | 2.58 | Range | 121 | Range | 1.13 |
| Minimum | 4000 | Minimum | 119 | Minimum | 2.98 | Minimum | 124 | Minimum | 2.89 |
| Maximum | 7661 | Maximum | 360 | Maximum | 5.56 | Maximum | 245 | Maximum | 4.02 |
| Sum | 251149 | Sum | 10600 | Sum | 209.79 | Sum | 8264 | Sum | 164.63 |
| Count | 50 | Count | 50 | Count | 50 | Count | 50 | Count | 50 |
| Confidence Level(95. | 255.015 | Confidence Level(95. | 16.3266 | Confidence Level(95. | 0.19008 | Confidence Level(95. | 8.74881 | Confidence Level(95 | 0.06012 |

Tab 22.19 Základné štatistické charakteristiky - 3. laktácia (zdroj: projekt KEGA, Candrák, 2021)

| *MLIEKO* | | *TUK* | | *TUKP* | | *BIELK* | | *BIELKP* | |
|---|---|---|---|---|---|---|---|---|---|
| Mean | 5231.81 | Mean | 214.702 | Mean | 4.10404 | Mean | 174.553 | Mean | 3.33426 |
| Standard Error | 124.857 | Standard Error | 6.30066 | Standard Error | 0.07651 | Standard Error | 4.62005 | Standard Error | 0.03436 |
| Median | 5180 | Median | 216 | Median | 4.12 | Median | 173 | Median | 3.32 |
| Mode | #N/A | Mode | 139 | Mode | 4.12 | Mode | 173 | Mode | 3.32 |
| Standard Deviation | 855.979 | Standard Deviation | 43.1952 | Standard Deviation | 0.52453 | Standard Deviation | 31.6735 | Standard Deviation | 0.23559 |
| Sample Variance | 732700 | Sample Variance | 1865.82 | Sample Variance | 0.27513 | Sample Variance | 1003.21 | Sample Variance | 0.0555 |
| Kurtosis | 1.34948 | Kurtosis | -0.1989 | Kurtosis | -0.5052 | Kurtosis | 0.73989 | Kurtosis | -0.5288 |
| Skewness | 0.97869 | Skewness | 0.2864 | Skewness | 0.19544 | Skewness | 0.84087 | Skewness | -0.0498 |
| Range | 3886 | Range | 171 | Range | 2.15 | Range | 136 | Range | 1.03 |
| Minimum | 4055 | Minimum | 136 | Minimum | 3.15 | Minimum | 121 | Minimum | 2.8 |
| Maximum | 7941 | Maximum | 307 | Maximum | 5.3 | Maximum | 257 | Maximum | 3.83 |
| Sum | 245895 | Sum | 10091 | Sum | 192.89 | Sum | 8204 | Sum | 156.71 |
| Count | 47 | Count | 47 | Count | 47 | Count | 47 | Count | 47 |
| Confidence Level(95. | 251.325 | Confidence Level(95. | 12.6826 | Confidence Level(95. | 0.15401 | Confidence Level(95. | 9.29968 | Confidence Level(95 | 0.06917 |

Výsledky sú samozrejme identické ako to bolo v prípade použitia programu SAS.

Výpočet základných štatistických charakteristík je možné uskutočniť v programe Excel aj pomocou vkladania jendotlivých štatistických funkcií. V tomto príklade by to bolo značne zložitejšie a časovo náročnejšie.

**Aktuálny prehľad základných štatistických funkcií programu Excel**

https://support.microsoft.com/en-us/office/statistical-functions-reference-624dac86-a375-4435-bc25-76d659719ffd

**Analýza údajov v programe Excel:**

https://support.microsoft.com/en-us/office/use-the-analysis-toolpak-to-perform-complex-data-analysis-6c67ccf0-f4a9-487c-8dec-bdb5a2cefab6

**Praktické použitie programu R** (R Studio)

**Príklad 1.1** (program R)

Vypočítajte a popíšte základné štatistické charakteristiky ukazovateľov mliekovej úžitkovosti kráv slovenského strakatého plemena. Určite, ktorý ukazovateľ má najväčšiu a najmenšiu variabilitu v hodnotenom súbore. Pre všetky hodnotené ukazovatele zostavte a interpretujte grafy vo forme histogramov a krabicových grafov. Zistite či hodnota obsahu tuku 4,00 % a hodnota obsahu bielkovín 3,30 % sa nachádza v rozpätí konfidenčných intervalov so spoľahlivosťou 95 %.

**Zadanie analýzy** (program R)

```
# Import a zobrazenie údajov (databáza projektu KEGA)
file_path <- "http://e-biostat.uniag.sk/wp-content/uploads/2022/01/Mlieko.txt"
Mlieko_R <- read.delim(file_path)
View(Mlieko_R)
# Inštalácia balíka summarytools (umožňuje veľkú rozmanitosť nastavení výpočtu)
install.packages("summarytools")
library(summarytools)
# Výpočet základných štatistických charakteristík
descr(Mlieko_R, digits=4)
descr(Mlieko_R,
  stats = "common") # most common descriptive statistics

# Rozšírený výpočet základných štatistických charakteristík
dfSummary(Mlieko_R, transpose = TRUE)

# Výpočet konfidenčných intervalov
install.packages("Rmisc")
library(Rmisc)
CI(Mlieko_R$TUKP, ci=0.95)
CI(Mlieko_R$BIELKP, ci=0.95)

# Zostavenie histogramov a krabicových grafov
hist(Mlieko_R$MLIEKO)
boxplot(Mlieko_R$MLIEKO ~ Mlieko_R$ROK)

boxplot(MLIEKO~PLEM,data=Mlieko_R, main="Produkcia mlieka",
    xlab="Plemenný typ", ylab="Mlieko kg")
# Výpočet základných štatistických charakteristík podľa triediaceho znaku
stby(data = Mlieko_R,
  INDICES = Mlieko_R$PL, # podľa poradia laktácie
  FUN = descr, # descriptive statistics
  stats = "common" ) # most common descr. stats
```

**Výsledky analýzy** (program R)

**Výpočet základných štatistických charakteristík**

N: 147

| | BIELK | BIELKP | MLIEKO | TUK | TUKP |
|---|---|---|---|---|---|
| Mean | 170.2517 | 3.3388 | 5099.8912 | 212.0680 | 4.1544 |
| Std.Dev | 30.1196 | 0.2221 | 844.0041 | 47.0321 | 0.5977 |
| Min | 121.0000 | 2.8000 | 4000.0000 | 119.0000 | 2.9700 |
| Q1 | 146.0000 | 3.1900 | 4431.0000 | 177.0000 | 3.7100 |
| Median | 168.0000 | 3.3100 | 4943.0000 | 210.0000 | 4.1200 |
| Q3 | 189.0000 | 3.4900 | 5610.0000 | 238.0000 | 4.5700 |
| Max | 257.0000 | 4.0200 | 7941.0000 | 360.0000 | 5.5600 |
| MAD | 31.1346 | 0.2224 | 839.1516 | 47.4432 | 0.6375 |
| IQR | 42.0000 | 0.2950 | 1153.0000 | 60.5000 | 0.8550 |
| CV | 0.1769 | **0.0665** | 0.1655 | **0.2218** | 0.1439 |
| Skewness | 0.7676 | 0.3969 | 0.9201 | 0.6133 | 0.3437 |
| SE.Skewness | 0.2000 | 0.2000 | 0.2000 | 0.2000 | 0.2000 |
| Kurtosis | 0.1377 | 0.1090 | 0.3967 | 0.2419 | -0.5154 |

Hrubo sú označené ukazovatele s najmenšou (BIELKP) a najväčšou variabilitou (TUK).

**Výpočet konfidenčných intervalov**

TUKP

| upper | mean | lower |
|---|---|---|
| 4.251849 | 4.154422 | 4.056994 |

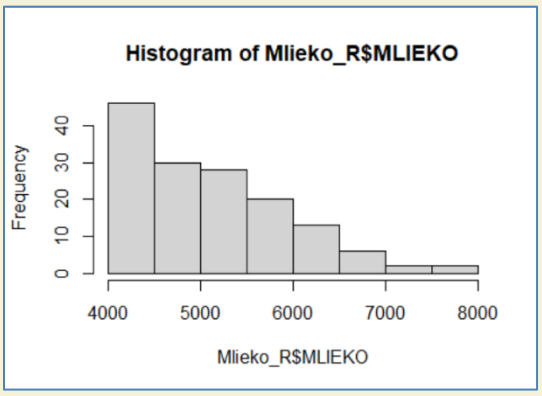
Hodnota 4.0 % ukazovateľa TUKP sa nenachádza v konfidenčnom intervale so spoľahlivosťou 95 %.

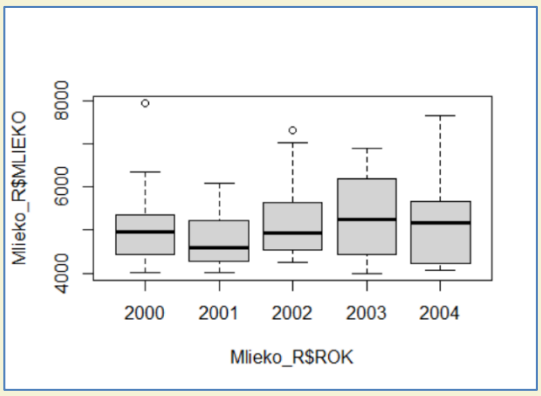
BIELKP

| upper | mean | lower |
|---|---|---|
| 3.375048 | 3.338844 | 3.302639 |

Hodnota 3.3 % ukazovateľa BIELKP sa nenachádza v konfidenčnom intervale so spoľahlivosťou 95 %.

Grafické zobrazenie (histogram a krabicový graf)



**Príklad 1.2** (program R)

Vypočítajte a popíšte základné štatistické charakteristiky ukazovateľa produkcia mlieka v kilogramoch podľa poradia laktácie. Zistite na ktorej laktácii má produkcia mlieka najväčšiu a najmenšiu variabilitu.

**Zadanie analýzy** (program R)

```
# Import a zobrazenie údajov (databáza projektu KEGA)
file_path <- "http://e-biostat.uniag.sk/wp-content/uploads/2022/01/Mlieko.txt"
Mlieko_R <- read.delim(file_path)
View(Mlieko_R)
# Inštalácia balíka summarytools (umožňuje veľkú rozmanitosť nastavení výpočtu)
install.packages("summarytools")
library(summarytools)
# Výpočet základných štatistických charakteristík podľa triediaceho znaku
stby(data = Mlieko_R,
  INDICES = Mlieko_R$PL, # podľa poradia laktácie
  FUN = descr, # descriptive statistics
  stats = "common") # most common descr. stats

# Inštalácia balíka psych
install.packages("psych")
library(psych)
# Výpočet základných štatistických charakteristík podľa triediaceho znaku
describeBy(
  Mlieko_R,
  Mlieko_R $PL) # grouping variable
```

**Výsledky analýzy** (program R)

**Výpočet základných štatistických charakteristík podľa poradia laktácie**

```
group: 01
 n   mean     sd median trimmed   mad  min  max range skew kurtosis     se
50 5052.8 777.48 4868.5 4960.73 617.5 4042 6891  2849 0.91    -0.23 109.95
group: 02
 n    mean     sd median trimmed    mad  min  max range skew kurtosis    se
50 5022.98 897.32 4788.5 4915.92 845.82 4000 7661  3661 0.91     0.04 126.9
group: 03
 n    mean     sd median trimmed    mad  min  max range skew kurtosis     se
47 5231.81 855.98   5180 5149.46 880.66 4055 7941  3886 0.92     0.91 124.86
```

**Výsledky analýzy** (program R)

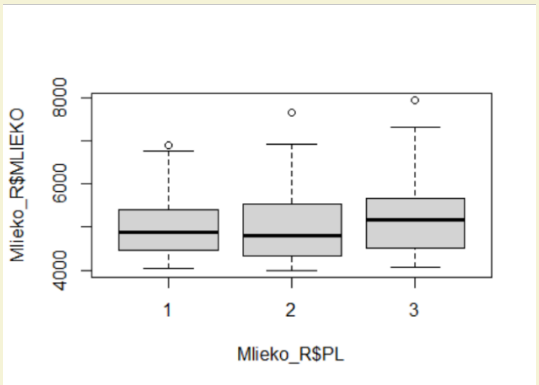
**Výpočet základných štatistických charakteristík podľa poradia laktácie**

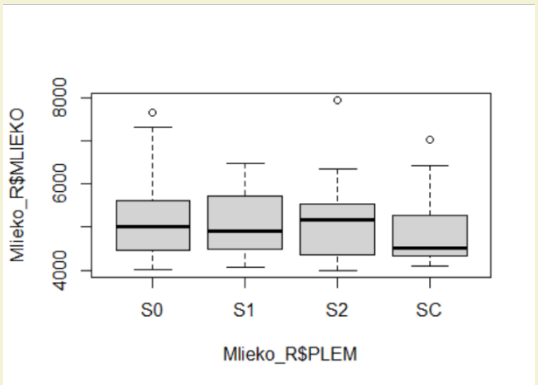
MLIEKO by PL

| | 1 | 2 | 3 |
|---|---|---|---|
| Mean | 5052.8000 | 5022.9800 | 5231.8085 |
| Std.Dev | 777.4833 | 897.3179 | 855.9789 |
| Min | 4042.0000 | 4000.0000 | 4055.0000 |
| Q1 | 4463.0000 | 4340.0000 | 4510.0000 |
| Median | 4868.5000 | 4788.5000 | 5180.0000 |
| Q3 | 5392.0000 | 5536.0000 | 5724.0000 |
| Max | 6891.0000 | 7661.0000 | 7941.0000 |
| MAD | 617.5029 | 845.8233 | 880.6644 |
| IQR | 925.2500 | 1183.5000 | 1159.0000 |
| CV | **0.1539** | **0.1786** | 0.1636 |
| Skewness | 0.9086 | 0.9081 | 0.9171 |
| SE.Skewness | 0.3366 | 0.3366 | 0.3466 |
| Kurtosis | -0.2251 | 0.0373 | 0.9131 |

Hrubo sú označené laktácie s najmenšou (1) a najväčšou variabilitou (2).

Grafické zobrazenie (krabicové grafy podľa poradia laktácie a plemenného typu)



**Zdroje a zoznam použitej literatúry**

Durkhure, P. and Lodwal, A., 2014. “Fault Diagnosis of Ball Bearing using Time Domain Analysis and Fast Fourier Transformation”. International Journal of Engineering Sciences & Research Technology, Vol. 3, pp.711-715

Zdroj 1: https://documentation.sas.com/doc/en/egdoccdc/8.3/egamotasks/titlepage.htm
Zdroj 2a: https://support.microsoft.com/en-us/office/statistical-functions-reference-624dac86-a375-4435-bc25-76d659719ffd
Zdroj 2b: https://support.microsoft.com/en-us/office/use-the-analysis-toolpak-to-perform-complex-data-analysis-6c67ccf0-f4a9-487c-8dec-bdb5a2cefab6
Zdroj 3: https://www.r-project.org/
The R Journal Vol. 13/1, June 2021 ISSN 2073-485.