

Projekt z předmětu Statistika

Téma: Typologie hráče české nejvyšší hokejové soutěže



VŠB-TU Ostrava:Fakulta Elektrotechniky a informatiky

jaro 2011

Martin Dočkal

doc068

dockal.martin@gmail.com

1 Obsah

2	Zadání	2
3	Úvod	3
4	Použitý Software	3
5	Explorační analýza	4
5.1	Ročník narození	4
5.2	Herní post	5
5.3	Odchovanec	7
5.4	Kanadské bodování.....	8
5.5	Postava (BMI).....	9
6	Metody statistické indukce	10
6.1	Odhad počtu hráčů v celé nejvyšší české hokejové lize, kteří letos oslaví své 18. narozeniny.....	10
6.2	Závislost věku hráče na jeho původu	11
6.3	Srovnání postavy hráčů a jejich herní pozice	14
6.4	Počet vstřelených branek u hráče v české nejvyšší hokejové lize.....	18
7	Závěr.....	19
8	Zdroje	20
9	Příloha A - datový soubor	20

2 Zadání

Zvolte si reálný výběrový soubor, který obsahuje alespoň

- 30 statistických jednotek
- 3 statistické proměnné

POZOR!!! Zpracovávány soubor musí být výběrovým souborem (vzorkem z nějaké populace). Pokud data nejsou náhodným výběrem, nelze je použít. Jednalo by se o tzv. vyčerpávající šetření, u něhož pozbývá smyslu celá statistická indukce.

Pro analýzu datového souboru použijte následující metody:

- Explorační analýza (povinně)

a alespoň jednu z každé skupiny uvedených metod statistické indukce:

- Intervalové odhady
- Jednovýběrové testy parametrických hypotéz
- Dvouvýběrové testy parametrických hypotéz
- ANOVA
- Analýza kontingenčních tabulek
- Regresní (jednoduchá lineární regrese) a korelační analýza

Součástí projektu je ověření všech předpokladů použitých metod statistické indukce.

3 Úvod

Účelem tohoto semestrálního projektu je prokázat schopnost správně demonstrovat a interpretovat datový záznam a schopnost provést v souladu s cílem projektu některé z metod statistické indukce, tedy aplikovat získané teoretické poznatky pomocí dostupného softwarového vybavení.

Zpracovávaným souborem jsem zvolil údaje hráčů české hokejové extraligy, konkrétně týmu HC Vítkovice a tyto údaje jsou v této práci vztažena k celé hokejové nejvyšší soutěži (výběr klubu, resp. hráčů zde tedy považuji za náhodný výběr).

Všechny informace o hráčích, resp. respondentech uvedených i zpracovávaných níže představuje sběr dat ze **základní části** soutěže ročníku **2010/11**.

4 Použitý Software

Microsoft Excel 2007

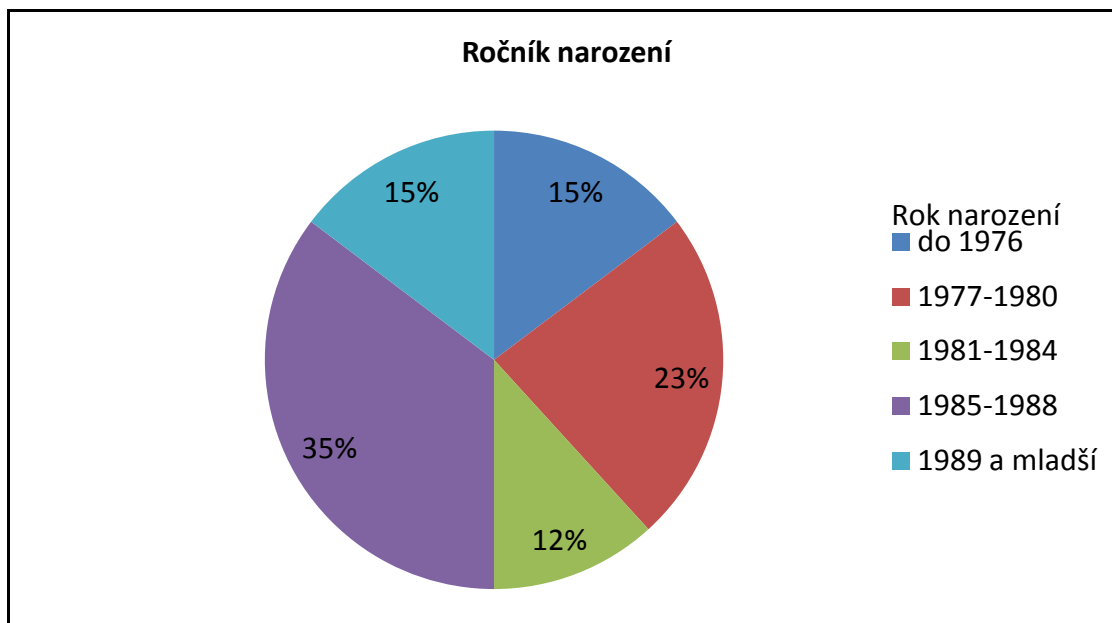
Statgraphics Plus 5.0.1.0

5 Explorační analýza

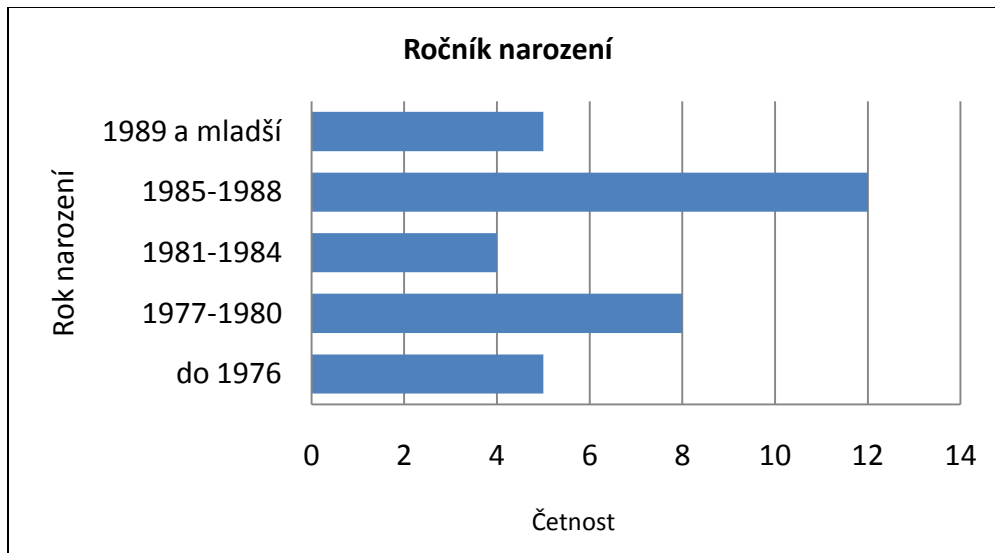
5.1 Ročník narození

Ročník narození				
Kategorie		Četnost		
ID	Hodnota	Absolutní	Relativní	%
1	do 1976	5	5/34	14,71
2	1977-1980	8	4/17	23,53
3	1981-1984	4	2/17	11,76
4	1985-1988	12	6/17	35,29
5	1989 a mladší	5	5/34	14,71
Celkem		34	1.0	100

Tabulka 1 : Rozdělení četností pro ročník narození



Graf 1: Výšečový graf pro ročník narození



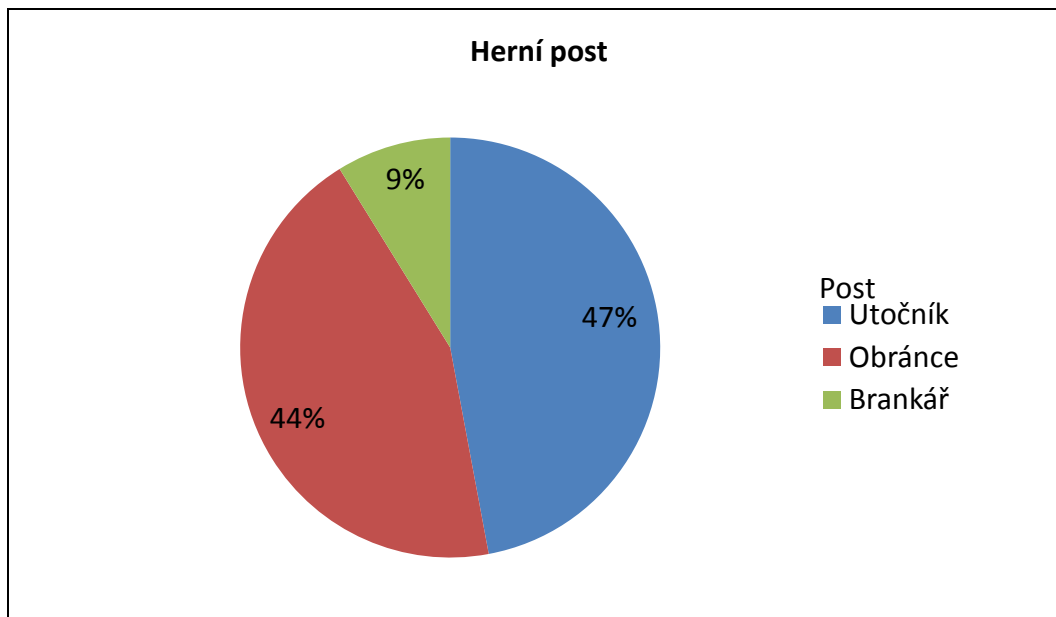
Graf 2 : Histogram pro ročník narození

Z výsledků našeho zkoumání je patrné, že jednotlivé věkové kategorie jsou v týmu poměrně rovnoměrně zastoupeny, výjimku tvoří skupina hráčů, kteří v letošním roce (2011) dovrší věku 23 až 26 let, která má v porovnání s ostatními nejčtenější zastoupení. V dnešním profesionálním sportu obecně platí snaha uplatnit a prosazovat mladé talentované hráče, aby mohli včas využít a rozvíjet jejich talent.

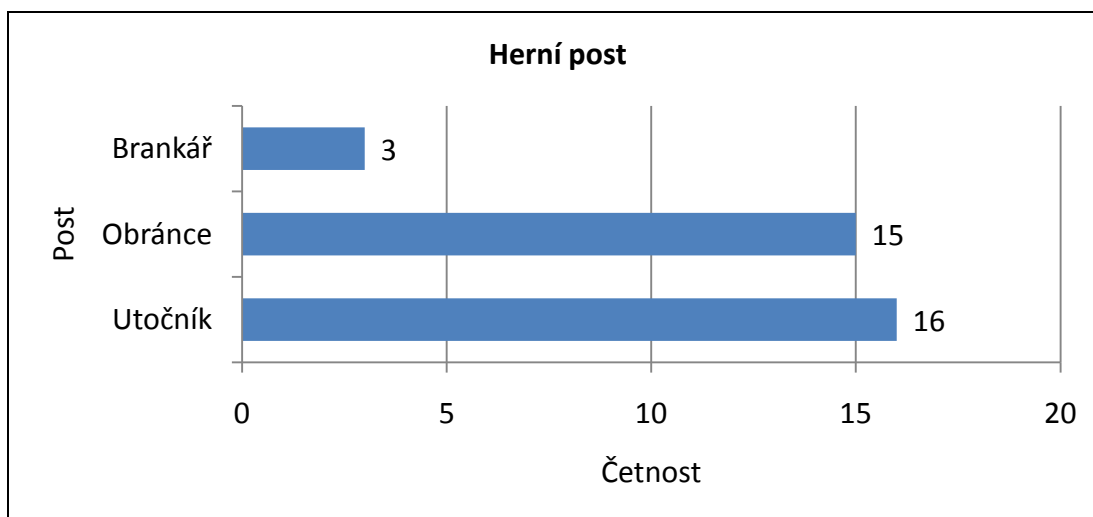
5.2 Herní post

Herní post				
Kategorie		Četnost		
ID	Hodnota	Absolutní	Relativní	%
1	Útočník	16	8/17	47,06
2	Obránce	15	15/34	44,12
3	Brankář	3	3/34	8,82
Celkem		34	1.0	100

Tabulka 2 : Rozdělení četností pro herní post



Graf 3 : Výšečový graf pro herní post



Graf 4 : Histogram pro herní post

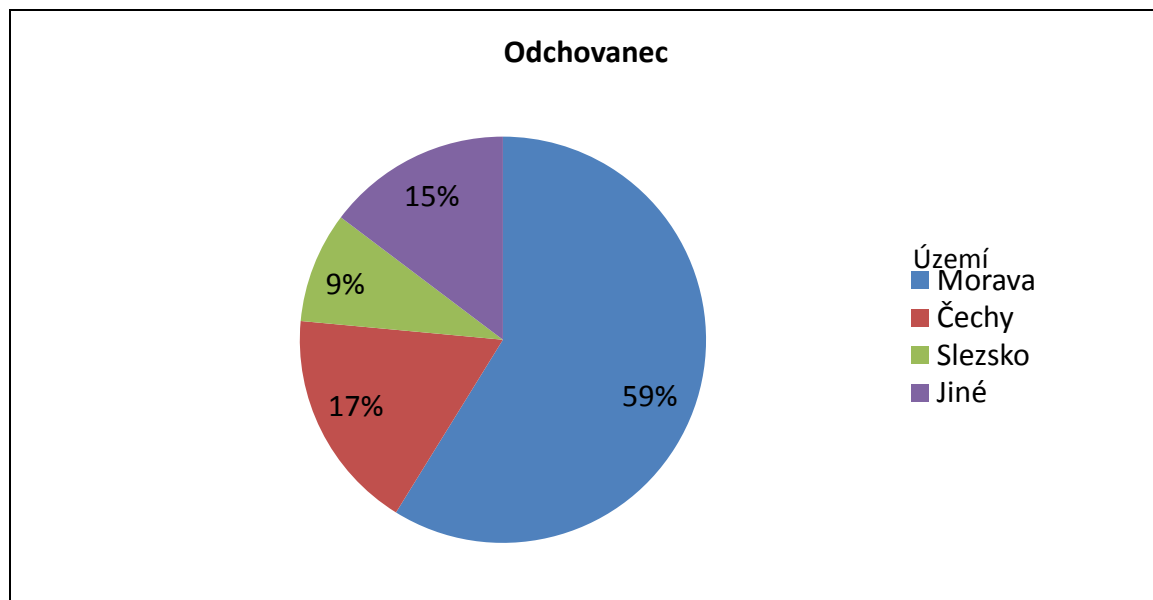
Každý hráč napsaný na soupisce pro herní sezónu 2010/11 má svou herní pozici. Ty existují právě tři, konkrétně brankář, obránce a útočník. Zatímco prvně jmenovaný obvykle nestřídá a odehraje celé utkání nepřetržitě, zbývající dvě role tvoří pětice hráčů v poměru 3:2 pro útočníky, jenž během zápasu v přibližně minutových intervalech střídají pro regeneraci sil. Není tedy žádné překvapení, že v námi zkoumaném týmu je nejčastější herní pozici právě útočník, překvapením ale je počet obránců, kteří do hry během sezóny zasáhli a jejich počet je jen o jednoho méně, nikoliv přibližně v poměru 2:3 z jejich pohledu, jak by bychom předpokládali. Znamená to tedy, že trenér měl větší potřebu obměňovat řady obránců buď z

vlastních mládežnických týmů, nebo výměnami s jinými kluby, ať už v případě nespokojenosti s herním projevem daného hráče či jeho zdravotním stavu.

5.3 Odchovanec

Odchovanec				
Kategorie		Četnost		
ID	Hodnota	Absolutní	Relativní	%
1	Morava	20	10/17	58,82
2	Čechy	6	3/17	17,65
3	Slezsko	3	3/34	8,82
4	Jiné	5	5/34	14,71
Celkem		34	1.0	100

Tabulka 3 : Rozdělení četností pro proměnnou odchovanec



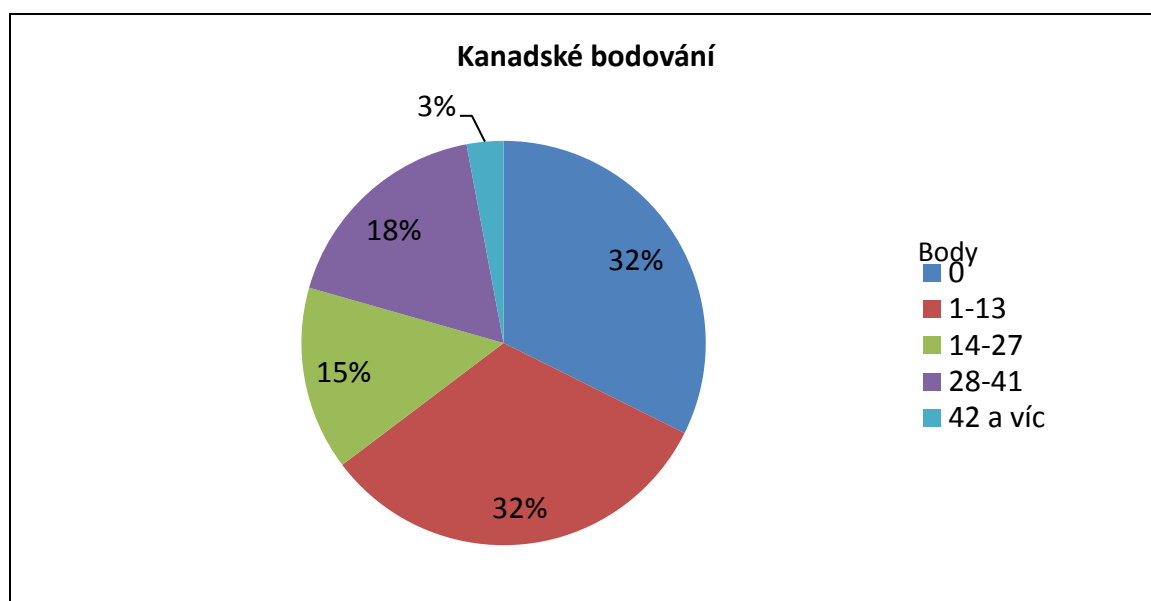
Graf 5 : Výšečový graf pro proměnnou odchovanec

Česká republika se dělí na 3 historické územní celky, ze kterých pochází 85% hráčů, zbytek tvoří cizinci. Nejvíce hráčů pak pochází z Moravy, což je z velké míry dáno umístěním města, ve kterém klub HC Vítkovice působí, tedy v tomto případě v Ostravě, jenž se nachází na rozhraní Moravy a Slezska.

5.4 Kanadské bodování

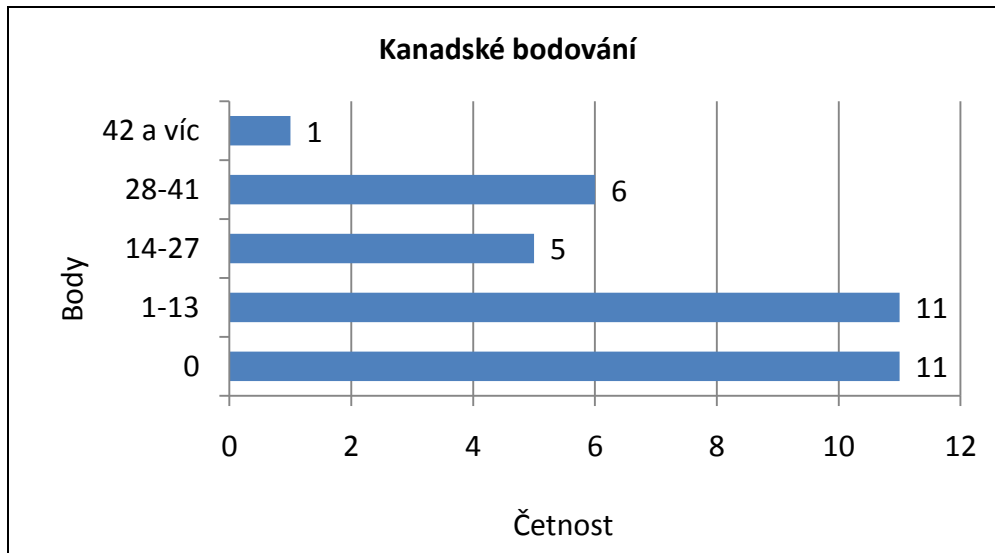
Kanadské bodování				
		Četnost		
ID	Hodnota	Absolutní	Relativní	%
1	0	11	11/34	32,35
2	1-13	11	11/34	32,35
3	14-27	5	5/34	14,71
4	28-41	6	2/17	17,65
5	42 a víc	1	1/17	2,94
Celkem		34	1.0	100

Tabulka 4 : Rozdělení četností pro kanadské bodování



Graf 6 : Výšečový graf pro kanadské bodování

Kanadské bodování je individuální statistika pro každého hráče, kterému se počítá za každou vstřelenou branku či nahrávku spoluhráči, který branku vsítil, jeden bod. Celkový součet je u každého hráče jednou z nejdůležitějších herních ukazatelů jeho výkonnosti vzhledem k herní produktivitě nejenom v rámci týmu, ale i celé soutěže. Z tabulky četností vidíme, že nejvíce hráčů nedosáhlo ani jediného bodu, případně jich nastřádali až celkem 13. První kategorii nám lehce zkreslili všichni brankáři, jenž si v dnešním moderním hokeji nepřipisují tolik bodů, jako v minulosti, protože ale je pořád teoreticky možné i regulérní, aby také brankář zaznamenal bod, do statistiky je započítat musíme.

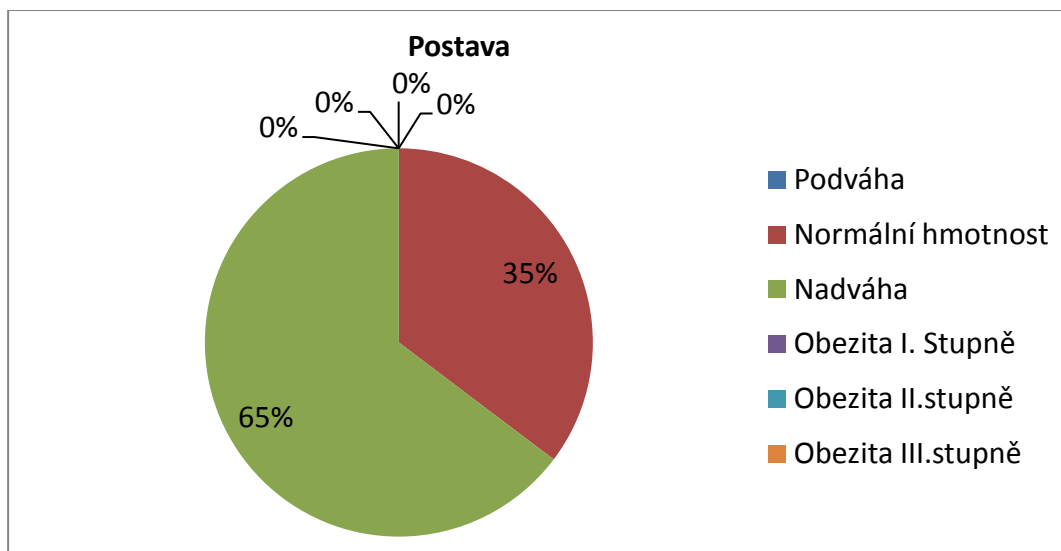


Graf 7 : Histogram pro kanadské bodování

5.5 Postava (BMI)

Postava				
Kategorie		Četnost		
ID	Hodnota	Absolutní	Relativní	%
1	Podváha	0	0/34	0,00
2	Normální hmotnost	12	6/17	35,29
3	Nadváha	22	11/17	64,71
4	Obezita I. Stupně	0	0/34	0,00
5	Obezita II.stupně	0	0/34	0,00
6	Obezita III.stupně	0	0/34	0,00
Celkem		34	1.0	100

Tabulka 5 : Rozdělení četností pro postavu (BMI)



Graf 8 : Výšečový graf pro postavu (BMI)

Poměr hmotnosti a výšky člověka nezávisle na jeho pohlaví nám udává tzv. index BMI, podle kterého jej pak řadíme do jedné z 6 definovaných kategorií. Přesný výpočet udává vztah

$$bmi = \frac{váha[kg]}{(výška[m])^2}$$

Jednotlivé kategorie pak pojmenováváme jako podváha (bmi do hodnoty 18,4), normální hmotnost (18,5 až 24,9), nadváha (25 až 29,9), obezita I. stupně (30 až 34,9), obezita II. stupně (35 až 39,9) a obezita III. stupně (40 a víc).

V našem případě se všichni respondenti, tedy hráči týmu HC Vítkovice rozdělili pouze do dvou skupin, z nichž nejpočetnější byla kategorie nadváhy. Jak ukazují poslední trendy ledního hokeje, který se řadí mezi nejrychlejší kolektivní hry na světě a patří mezi fyzické sporty, mírný tělesný nadbytek nebývá v tomto případě na škodu, navíc je nutné brát na zřetel fakt, že u dotazovaných profesionálních hráčů je předpoklad většího objemu svalové hmoty, než u ostatních byt' rekreačně sportujících lidí, která bezesporu zvyšuje jejich hmotnost.

6 Metody statistické indukce

6.1 Odhad počtu hráčů v celé nejvyšší české hokejové lize, kteří letos oslaví své 18. narozeniny

Využijeme skutečnosti, že od všech jednotlivých hráčů vybraného týmu, resp. respondentů náhodného týmu máme přesně daný rok narození a na základě tohoto faktu a metody **intervalových odhadů** se pokusíme odhadnout počet hráčů v celé české hokejové nejvyšší lize, kterou dnes čítá celkem dalších 13 klubů, s rokem narození **1993**, tedy takových, kteří v letošním roce (2011) oslaví své **18. narozeniny**. Tento údaj tedy budeme považovat za náhodnou veličinu s normální rozdělením. Pro pořádek uvedme množinu našich získaných dat, se kterou budeme pracovat, jenž je shodná s datovou přílohou na konci dokumentu:

$$M = \left\{ \begin{array}{l} 1977,1986,1980,1972,1985,1975,1990,1986,1989,1980,1986,1983,1987,1984, \\ 1976,1976,1984,1989, 1985,1989,1990,1991,1977,1981,1987,1977, \\ 1987,1992,1991,1980,1988,1977,1979,1975 \end{array} \right\}$$

Máme celkem 34 hodnot, tedy $n = 34$.

Průměr μ je roven vztahu $\mu = \frac{\sum_{k=1}^n M_k}{n} = \frac{67431}{34} \doteq 1983,26$.

Směrodatná odchylka $\delta \doteq 5,67$.

95% interval spolehlivosti (confidence interval) pro průměr $\mu \doteq 1983,26$ je roven (1981,28; 1985,24).

Výpis z programu Statgraphics pro tento úkon:

```
Sample mean = 1983,26
Sample standard deviation = 5,67
Sample size = 34

95,0% confidence interval for mean: 1983,26 +/- 1,97836 [1981,28;1985,24]
```

Obrázek 1 : Určení intervalu spolehlivosti pro průměr v případě intervalových odhadů

95% interval spolehlivosti (confidence interval) pro odchylku $\delta = 5,67$ je roven (4,57; 7,46). Výpis z programu Statgraphics pro tento úkon:

```
Sample standard deviation = 5,67
Sample size = 34

95,0% confidence interval for sigma: [4,57329;7,46329]
```

Obrázek 2 : Určení intervalu spolehlivosti pro odchylku v případě intervalových odhadů

$$X \rightarrow N(1981,28; 4,57^2) \Rightarrow P(X = 1993) \doteq 0,003$$

$$X \rightarrow N(1981,28; 7,46^2) \Rightarrow P(X = 1993) \doteq 0,016$$

$$X \rightarrow N(1985,24; 4,57^2) \Rightarrow P(X = 1993) \doteq 0,021$$

$$X \rightarrow N(1985,24; 7,46^2) \Rightarrow P(X = 1993) \doteq 0,031$$

\Rightarrow na základě intervalových odhadů nad daty o výběru 34 hráčů z klubu, který považujeme náhodný, odhadujeme, že v české nejvyšší lize v letošním roce oslaví své 18. narozeniny s 95% spolehlivostí 0,3% až 3,1% hráčů.

6.2 Závislost věku hráče na jeho původu

Budeme zkoumat, zda věk hráče v týmu souvisí s jeho místem počátku působení v ledním hokeji (odchovanec) v rámci nejvyšší soutěže. Využijeme k tomu tzv. **kontingenčních tabulek**.

Nezávislou proměnnou zde tedy máme historickou část České republiky (Morava, Slezsko, Čechy) případně jinou zemi (např. Slovensko).

Stanovení hypotéz:

Nulová hypotéza H_0 : Věk hráče **nezávisí** na místě jeho původu.

Alternativní hypotéza H_A : Věk hráče **závisí** na místě jeho původu.

K výsledku se dopracujeme sérií tabulek marginálních četností.

Frequency Table

	do 1976	od 77 do 80	od 81 do 84	od 85 do 88	nad 1988	Row Total
Morava	3 8,82%	2 5,88%	3 8,82%	6 17,65%	6 17,65%	20 58,82%
Čechy	1 2,94%	4 11,76%	0 0,00%	1 2,94%	0 0,00%	6 17,65%
Slezsko	0 0,00%	0 0,00%	0 0,00%	1 2,94%	2 5,88%	3 8,82%
Ostatní	1 2,94%	2 5,88%	1 2,94%	1 2,94%	0 0,00%	5 14,71%
Column Total	5 14,71%	8 23,53%	4 11,76%	9 26,47%	8 23,53%	34 100,00%

Cell contents:
Observed frequency
Percentage of table

Obrázek 3 : Počáteční tabulka marginálních četností pro metodu kontingenčních tabulek

Z první, a tedy počáteční tabulky vidíme, že mnohé četnosti klesají pod přípustnou hodnotu (2) a ani většina buněk (alespoň 80%) nemá četnost alespoň 5 a protože tak nejsou splněny předpoklady pro relevantní určení hodnoty p-value za účelem zamítnutí nulové, resp. alternativní hypotézy metodou kontingenčních tabulek, nelze v metodě zatím pokračovat. Sloučíme tedy řádky č. 2, 3 a 4.

Frequency Table

	do 1976	od 77 do 80	od 81 do 84	od 85 do 88	nad 1988	Row Total
Morava	3 8,82%	2 5,88%	3 8,82%	6 17,65%	6 17,65%	20 58,82%
Ostatní	2 5,88%	6 17,65%	1 2,94%	3 8,82%	2 5,88%	14 41,18%
Column Total	5 14,71%	8 23,53%	4 11,76%	9 26,47%	8 23,53%	34 100,00%

Cell contents:
Observed frequency
Percentage of table

Obrázek 4 : Tabulka marginálních četností pro metodu kontingenčních tabulek po prvním sloučení

Opět nejsou předpoklady splněny, sloučíme tedy sloupce 1 a 2, dále také 3 a 4.

Frequency Table

	do 1980	od 81 do 88	nad 1988	Row Total
Morava	5 14,71%	9 26,47%	6 17,65%	20 58,82%
Ostatní	8 23,53%	4 11,76%	2 5,88%	14 41,18%
Column Total	13 38,24%	13 38,24%	8 23,53%	34 100,00%

Cell contents:

Observed frequency
Percentage of table

Obrázek 5 : Tabulka marginálních četností pro metodu kontingenčních tabulek po druhém sloučení

Nyní už máme většinu hodnot buněk alespoň 5 a žádná neklesla pod 2. Přejdeme tedy k určení hodnoty p -value a zamítnutí nebo nezamítnutí nulové hypotézy, resp. alternativní hypotézy určených v úvodu této podkapitoly.

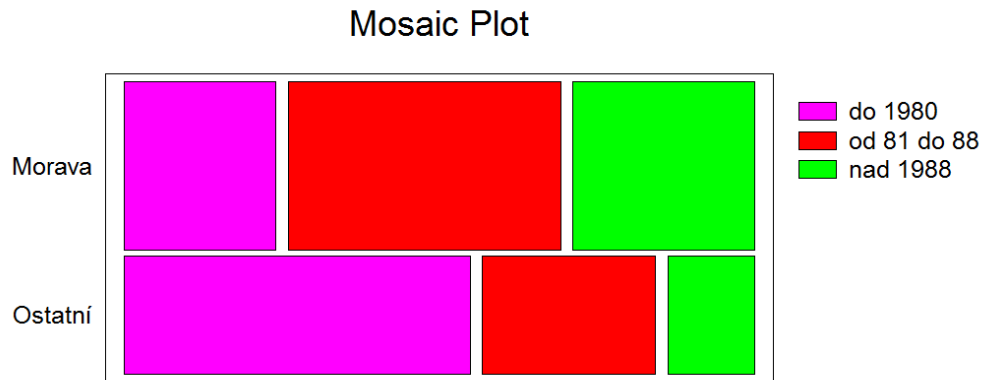
Chi-Square Test

Chi-Square	Df	P-Value
3,67	2	0,1595

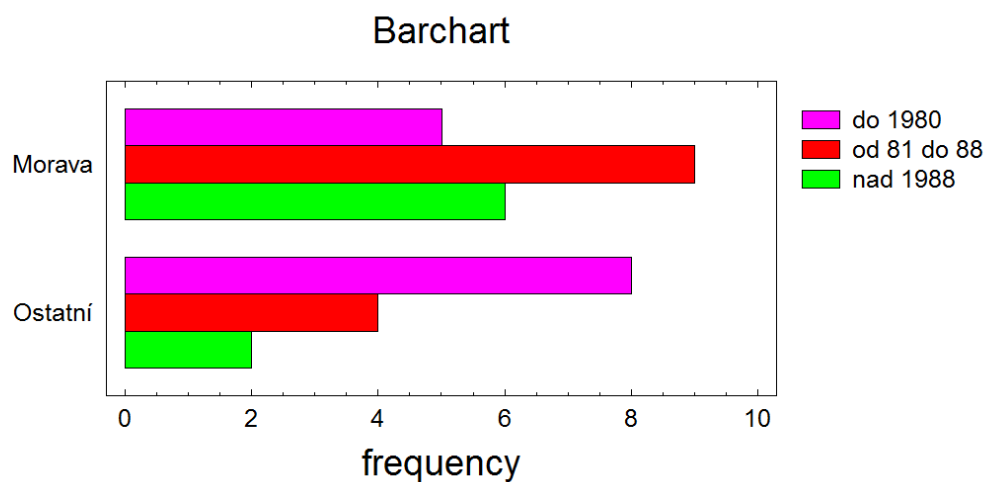
Warning: some cell counts < 5.

Obrázek 6 : Určení hodnoty p -value (X-kvadrát testu)

Protože hodnota p – value > 0,05, zamítáme na hladině významnosti 0,05 alternativní hypotézu ve prospěch nulové hypotézy, resp. nezamítáme nulovou hypotézu. Věk hráče tedy **nezávisí** na místě jeho původu.



Graf 9 : Mozaikový graf v případě určování závislosti věku hráče na místě jeho původu



Graf 10 : Histogram v případě určování závislosti věku hráče na místě jeho původu

Výsledek není vzhledem k menší územní rozloze nijak překvapující. Svůj vliv zajisté měl fakt, že klub se nachází na rozmezí dvou historických územních celků a většina našich respondentů tak mohla upřednostnit spíše subjektivně či nevědomky jednu variantu před druhou.

Pro úplnost dodáme, že stejného výsledku, tedy nezamítnutí nulové hypotézy bychom dosáhli také v případě, že bychom striktně trvali na alespoň 80% buněk a nestačila by většina, u kterých by hodnota četnosti byla rovna alespoň hodnotě 5 a dále poslední marginální tabulku slučovali.

6.3 Srovnání postavy hráčů a jejich herní pozice

Budeme analyzovat, zda v české hokejové nejvyšší soutěži platí nějaký vztah mezi postavou, konkrétně jeho absolutní hmotností a typem postavy (index BMI, viz výše). Herní pozicí

myslíme kategoriální proměnnou Herní pozice nabývající právě hodnot útočník, obránce, nebo brankář. Využijeme metody **ANOVA**.

Stanovení první sady hypotéz:

Nulová hypotéza H_0 : Hmotnost hráče **nezávisí** na jeho herní pozici.

Alternativní hypotéza H_A : Hmotnost hráče **závisí** na místě jeho původu.

Shrneme si tabulkou pro jednotlivé druhy herních pozic nasbíraná data, v tomto první případě se jedná o hmotnosti hráčů v kg:

útočník	obránce	brankář
83	105	76
86	86	81
81	102	86
88	95	
70	90	
81	90	
86	90	
84	88	
92	79	
92	100	
85	85	
96	85	
76	94	
80	95	
85	96	
70		

Tabulka 6 : Výčet hmotností hráčů v kg pro jednotlivé herní role

Nejprve jsme museli potvrdit normalitu dat každé třídy (v záhlaví tabulky) jako nezbytný předpoklad pro metodu ANOVA. Ověřovali jsme nulovou hypotézu, zda všechny tři výběry lze považovat za náhodný výběr z normálního rozdělení (pro každou třídu zvlášť).

Ve všech třech případech nám vyšla hodnota $p - value > 0,05$, díky čemuž nezamítáme nulovou hypotézu a výběry tak považujeme za náhodný výběr z normálního rozdělení.

Obdobným způsobem jsme testovali podmínku tzv. homoskedasticity, tj. neexistence statisticky významného rozdílu mezi rozptyly jednotlivých tříd (nulová hypotéza). Také zde jsme pro jednotlivé výběry dosáhli hodnoty $p - value > 0,05$, potvrdili tak nulovou hypotézu a splnili další předpoklad pro analýzu metodou ANOVA.

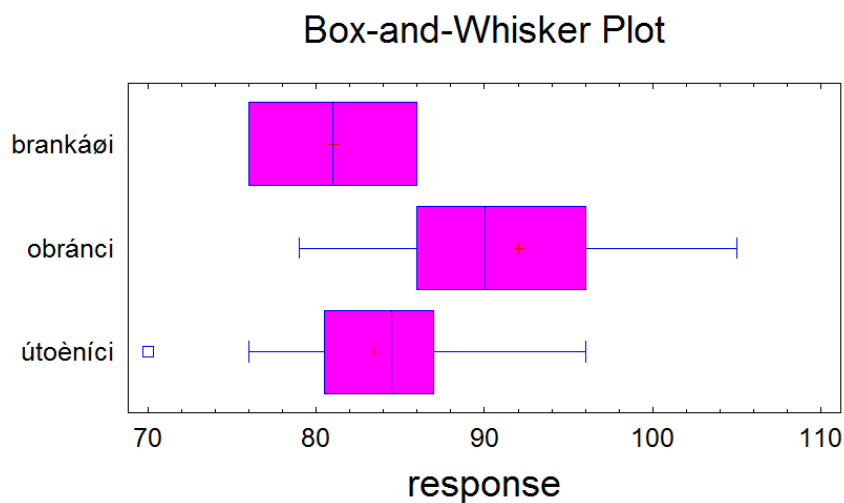
Poslední podmínkou je nezávislost výběrů, jenž je zřejmá ze samotného zadání.

Přejdeme tedy k určení p – *value* pro zamítnutí nulové, nebo alternativní hypotézy ve prospěch své opačné alternativy.

ANOVA Table

Analysis of Variance					
Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Between groups	686,063	2	343,031	6,92	0,0033
Within groups	1535,94	31	49,5464		
Total (Corr.)	2222,0	33			

Obrázek 7 : Určení p -value za účelem rozhodnutí v ANOVA metodě pro hmotnost hráče s ohledem na herní pozici



Graf 11 : Vícenásobný krabicový graf pro metodu ANOVA a určení souvislosti hmotnosti hráče a jeho herní pozice

Protože je p – *value* $< 0,05$, zamítáme H_0 ve prospěch H_A , resp. nulovou hypotézu ve prospěch její alternativy. Na základě těchto údajů byla potvrzena závislost mezi absolutní hmotností hráče a jeho herní pozicí.

Pro úplnost si zavedeme ještě další hypotézu s ohledem na výšku hráče vzhledem k jeho hmotnosti (index BMI), neboť hráč může mít vysokou hmotnost i vysokou výšku (a naopak).

Nulová hypotéza H_0 : BMI hráče **nezávisí** na jeho herní pozici.

Alternativní hypotéza H_A : BMI hráče **závisí** na místě jeho původu.

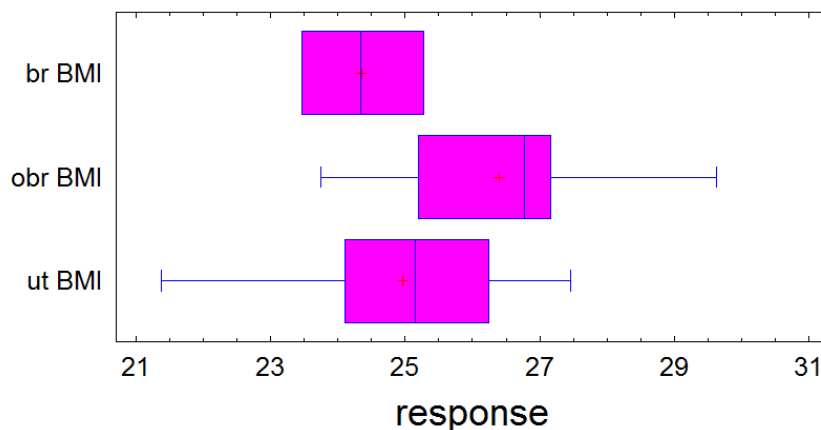
Shrneme si tabulkou pro jednotlivé druhy herních pozic nasbíraná data, v tomto druhém případě se jedná o hodnoty indexu BMI hráčů (bezrozměrná veličina, resp. $kg \cdot m^{-2}$):

útočník	obránce	brankář
25,05736	26,78298	23,45679

26,25072	26,25072	25,28011
25,56495	27,66927	24,33228
27,46481	27,16692	
22,59814	26,8745	
26,44898	25,46401	
24,59321	25,19526	
24,54346	27,16049	
26,59267	24,65591	
25,48476	27,12674	
26,23457	24,30724	
25,24655	26,23457	
23,71961	23,73677	
24,41928	27,75749	
23,79553	29,62963	
21,36687		

Tabulka 7 : Výčet hmotností hráčů v kg pro jednotlivé herní role

Box-and-Whisker Plot



Graf 12 : Vícenásobný krabicový graf pro metodu ANOVA a určení souvislosti indexu BMI hráče a jeho herní pozice

Také zde jsme testovali, obdobně jako v minulém případě pro hmotnosti hráčů u herních pozic (tříd) předpoklady pro analýzu rozptylu metody ANOVA, kdy i v tomto případě pro indexy BMI jsme u všech výběrů splnili jak podmínku považování těchto výběrů za náhodný výběr s normálním rozdělením, testu homoskedasticity i nezávislosti výběrů (patrně opět ze zadání) a přejdeme rovnou ke stanovení hodnoty p-value a rozhodnutí o zamítnutí hypotéz:

ANOVA Table

Analysis of Variance					
Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Between groups	20,6998	2	10,3499	4,48	0,0195
Within groups	71,5419	31	2,3078		
Total (Corr.)	92,2417	33			

Obrázek 8 : Určení p-value za účelem rozhodnutí v ANOVA metodě pro BMI hráče s ohledem na herní pozici

Také v tomto případě, kdy jsme vedle samostatné hmotnosti hráče zohlednili i jeho výšku, jsme dospěli k hodnotě $p - value < 0,05$. Zamítáme tedy nulovou hypotézu ve prospěch alternativy a tvrdíme, že na základě těchto údajů na hladině významnosti 0,05 byla potvrzena závislost mezi postavou hráče v podobě jeho indexu BMI a jeho herní pozice.

Výsledky naznačují, že u obránců je kladen větší důraz na jeho fyzické proporce, což vyplývá z podstaty hry určenou pro tuto roli na hřišti.

6.4 Počet vstřelených branek u hráče v české nejvyšší hokejové lize

Opět využijeme našeho datového souboru a skutečnosti, že máme počty vstřelených branek u hráčů, jehož vzorek tak považujeme za náhodný výběr vztažený k celé naší české hokejové extralize. Uvedeme množinu dat (vstřelených branek) u každého našeho respondenta:

$$M = \{15,19,18,17,15,20,5,8,5,4,7,5,1,3,2,0,5,4,4,2,1,3,1,0\}$$

Pro určení zda hráč vstřelil v uplynulé sezóně 2010/11 alespoň jednu branku, použijeme metodu **Jednovýběrový test parametrických hypotéz**.

Stanovení hypotéz:

Nulová hypotéza H_0 : Hráč v extralize nevstřelil žádnou branku.

Alternativní hypotéza H_A : Hráč v extralize vstřelil v průměru alespoň jednu branku.

Hypothesis Tests

Sample mean = 4,82
 Sample standard deviation = 6,33
 Sample size = 34

99,0% lower confidence bound for mean: 4,82 - 2,65404 [2,16596]

Null Hypothesis: mean = 0,0
 Alternative: greater than
 Computed t statistic = 4,44
 P-Value = 0,0000476414
 Reject the null hypothesis for alpha = 0,01.

The StatAdvisor

This analysis shows the results of performing a hypothesis test concerning the mean (μ) of a normal distribution. The two hypotheses to be tested are:

Null hypothesis: $\mu = 0,0$
 Alternative hypothesis: $\mu > 0,0$

Given a sample of 34 observations with a mean of 4,82 and a standard deviation of 6,33, the computed t statistic equals 4,44. Since the P-value for the test is less than 0,01, the null hypothesis is rejected at the 99,0% confidence level. The confidence bound shows that the values of μ supported by the data are greater than or equal to 2,16596.

Obrázek 9 : Souhrn výsledků pro hypotézu

Protože hodnota $p - value < 0,01$, zamítáme nulovou hypotézu ve prospěch alternativy, tedy tvrdíme, na hladině významnosti 0,01, že hráč v extralize dosáhl střední hodnoty počtu vstřelených branek větších než 0, tedy vstřelil alespoň jednu branku.

7 Závěr

Zpracováním a analýzou vzorku dat nad vybraným hokejovým klubem, resp. hráči jsme dospěli např. mj. k následujícím výsledkům. vztaheným k celé hokejové soutěži:

Na soupisce se vyskytuje téměř rovnovážný počet obránců i útočníků

Až 3,1 % hráčů oslaví nebo již oslavilo v letošním roce (2011) své 18. narozeniny.

Věk hráče nezávisí na místě jeho původu.

Postava hráče, jak v podobě jeho absolutní váhy, tak i v poměru k výšce (index BMI) je závislá na jeho herní pozici v týmu a na hřišti (obránce, útočník, brankář).

Hráč nastřílel alespoň jednu branku.

8 Zdroje

[1] **Šimonová, Lenka.** *Průvodce k programu Statgraphics.* Ostrava : Vysoká škola báňská - Technická univerzita Ostrava, Fakulta elektrotechniky a informatiky, Katedra aplikované matematiky, 2006.

[2] **Litschmannová, Martina Ing.** *Slajdy ze cvičení Statistika I.*

[3] **Janurová, Kateřina Ing.** *Cvičení na téma ANOVA.* Ostrava : autor neznámý, 2011.

[4] **Vítkovice, HC.** Statistika hráčů. *Oficiální internetové stránky klubu HC Vítkovice.* [Online] [Citace: 16. Duben 2011.] www.hc-vitkovice.cz.

9 Příloha A - datový soubor

viz další list