

Česká zemědělská univerzita v Praze

Provozně ekonomická fakulta

Katedra statistiky



Teze bakalářské práce

**Statistická analýza faktorů ovlivňujících výkon na
vybraných postech v softbalu**

Tomáš Klein

© 2015 ČZU v Praze

Souhrn

Tato bakalářská práce se zabývá vyhodnocením průměrných výkonů vybraných hráčů softballu v České republice v letech 2010 – 2014 a jejich následnou statistickou analýzou.

Práce se dělí na tři části. V první části je popsána metodika týkající se parametrických testů, analýzy rozptylu a regresní analýzy. Ve druhé části jsou teoretické údaje o sportovním výkonu a jeho faktorech. Poslední částí je praktická část, kde je zpracována samotná analýza parametrických testů, analýzy rozptylu o průměrných výkonech catcherů a ostatních hráčů. A závislost výkonu pálkařského průměru na počtu startů na pálce.

Klíčová slova: softball, sportovní výkon, faktory sportovního výkonu, popisná statistika, testování hypotéz, regresní analýza

1. Cíl práce a metodika

Cílem této práce je vyhodnocení výkonů vybraných postů v softballu. Jako dílčí cíle jsou stanoveny dvě pracovní hypotézy, které budeme ověřovat: 1) Catcheři dosahují v průměru lepších výkonů než hráči na ostatních postech a za 2) Výkonnost je závislá na počtu startů na pálce. Předpokladem pro tuto práci je domněnka trenérů, že hráči na specifickém postu catchera (chytáče) jsou lepšími pálkaři než ostatní hráči na jiných postech. Vzhledem k tomu, že v obraně neustále chytají nadhozy svým nadhazovačům, tak častěji sledují umístění míčku do strike zóny, dráhu míčku a rychlost míčku. Tyto všechny faktory by jim měly dát teoreticky výhodu lepších pálkařů, neboť by měli lépe číst nadhozené míče. Proto vznikl předpoklad trenérů, že catcherové by měli být obecně lepšími pálkaři.

Z dat útočných statistik se vytvořily dva nezávislé výběrové soubory a to soubor catcherů a ostatních hráčů. Z těchto souborů se vždy separovalo pouze 25% nejvytíženějších hráčů. První hypotéza se zkoumala parametrickým testem. Vybranými statistickými znaky byly BA (pálkařský průměr), OBA (úspěšnost dosažení první mety) a SA (průměr met dosažených odpalem). Na základě p-hodnoty t-testu se následně na hladině významnosti $\alpha = 0,05$ zamítla či potvrdila nulová hypotéza o tom, zda jsou catcheři lepšími pálkaři. Analýza rozptylu zkoumá vliv jednoho či více faktorů na zkoumaný znak. Testovala se každá skupina zvlášť (catcherové, ostatní hráči, všichni hráči dohromady). Jako faktor u této metody figuroval rok. Zkoumaným znakem byl pálkařský průměr. Metoda zkoumá variability proměnných. Dělí se na meziskupinovou a vnitroskupinovou. Měření probíhá podle součtu čtverců. Testovým kritériem je poměr rozptylu. A ten vychází z nulové hypotézy, že skupinové průměry nejsou statisticky významné. K ověření předpokladu se používal Bartlettův test.

Přímková regrese vyjadřuje změny jedné veličiny při změně jiné veličiny. Nejprve se stanovily proměnné (nezávislá, závislá). Nejčastějším typem regresní funkce je přímka. Pro přímku se stanovují parametry. Ty se stanoví metodou nejmenších čtverců dosazenou do regresní přímky. Koeficientem korelace se měří těsnost závislosti pro libovolnou regresní funkci, jejíž parametry se odhadly pomocí metody nejmenších čtverců. Koeficient determinace je druhou mocninou koeficientu korelace. Představuje míru těsnosti lineární závislosti. Udává, z kolika procent jsou změny závisle proměnné vysvětlitelné zvolenou lineární funkcí. K vypracování bakalářské práce posloužila řada odborných publikací věnujících se dané problematice. Podkladová data poskytly internetové stránky České softballové asociace. Pro jednotlivé výpočty posloužil software Statistica 12.

2. Výsledky a závěr

Test průměrného výkonu BA, OBA, SA

Pálkařský průměr BA vypovídá pouze o výkonu samotného hráče. Statistika OBA je obsáhlejší než pálkařský průměr. A vypovídá o hráčově úspěšnosti obsadit první metu. Zahrnuje nejen samotný hráčův výkon, ale také chyby, kterých se dopustil soupeř v hráčův prospěch. Statistika SA (průměr met dosažených odpalem). Tento ukazatel nám prokazoval, zda v průměru catcherové dokážou na jeden start na pálce napálit více met než zbytek hráčů ligy. V těchto testech jsme porovnávali skupinu catcherů a ostatních hráčů, zda se průkazně liší výkon těchto pálkařského znaků u catcherů oproti ostatním hráčům ligy. Neboli zda existuje statisticky významný rozdíl mezi catchery a ostatními hráči. Výsledky p-hodnoty t-testu poukázaly, že ani v jednom ze zmiňovaných znaků v období za roky 2010 – 2014 na hladině významnosti $\alpha = 0,05$ se hypotéza, že by byli catcherové lepšími pálkaři a dosahovali tak lepších výkonů nepotvrdila.

ANOVA

Tuto metodu jsme použili, abychom otestovali, zda je průkazný rozdíl v průměrném výkonu BA mezi jednotlivými roky u catcherů a ostatních hráčů. V této metodě se zformulovaly statistické hypotézy pro tři analýzy. První analýza byla, zda je rozdíl v průměrném výkonu pálkařského průměru catcherů za jednotlivé roky. Druhá analýza byla, zda je rozdíl v průměrném výkonu pálkařského průměru ostatních hráčů za jednotlivé roky. A třetí analýza, zda je rozdíl v průměrném výkonu všech hráčů dohromady za jednotlivé roky. Ve všech třech analýzách, kde jako faktor figuroval rok, byla vždy p-hodnota F-testu $>$ než 0,05 a tedy v žádném z těchto případů se neprokázal statisticky významný rozdíl.

Závislost výkonu BA na počtu startů na pálce catcherů

V této analytické části jsme chtěli zjistit, zda existuje závislost mezi výkonem BA na počtu startů na pálce (AB). Jako závislá proměnná figuroval výkon BA a nezávislá proměnná AB. Odhadnuté parametry přímky vyšly $a = 0,182$ a $b = 0,001$. Regresní funkce měla tedy tvar $y' = 0,182 + 0,001x$. Na základě p-hodnoty 0,078 jsme mohli stanovit, že oba parametry a i b jsou na hladině významnosti $\alpha = 0,05$ statisticky nevýznamné. Korelační koeficient byl $r = 0,306$ který nám říká, že lineární závislost je slabá. Hodnota koeficientu determinace $r^2 = 0,093$ nám udává, že z 9,3% jsou změny ve výkonu BA ovlivněny změnami účasti na pálce (AB). Nepodařilo se nám nalézt regresní model, který by tuto závislost vysvětlil. Na druhou stranu

to znamená, že máme vysoké procento (přes 90%) variability dat v závisle proměnné, která není modelem vysvětlena. Na základě výsledku t-testu, kterým se testoval regresní koeficient, a p-hodnoty můžeme prohlásit, že H_0 : o nezávislosti nelze zamítnout. Výsledky tohoto testu tedy nelze zobecnit.

Závěr

Bakalářská práce se zabývala vyhodnocením průměrných výkonů na pálce u vybraného postu catchera v porovnání s ostatními nejlepšími ligovými hráči na pálce za období let 2010 - 2014. Obecně se mělo za to, že catcherové ze své pozice lépe vnímají prostor strike zóny. Vyhodnocování bylo testováno pomocí statistických analýz. Tyto analýzy nám odpověděly na zmiňovaný předpoklad trenérů. Na základě získaných dat z České softballové asociace a jejich následnou separací se porovnávali dvě skupiny nejvytíženějších hráčů. Toto porovnání bylo provedeno pomocí statistických testů. Konkrétně se testoval pálkařský průměr, procentuální dosažení první mety a průměr napálených met na jeden start na pálce. V případě první hypotézy se v žádném z případů nepotvrdil statisticky významný rozdíl mezi catchery a ostatními hráči. Na základě zpracování parametrických testů se tato trenérská domněnka nepotvrdila. Catcherové se jako významně lepší pálkaři neprojevovali. V případě druhé hypotézy o závislosti pálkařského průměru na počtu startů na pálce tak ani zde se nic statisticky významného nepotvrdilo. Výkon pálkařů se nezlepšuje podle absolutních počtů startů na pálce, ale podle komplexních zkušeností hráče, individuálním citem pro hru, jenž hráč získává během své sportovní kariéry.

Jako doporučení softballovým trenérům lze stanovit jednoznačně to, že když vybírají hráče do zápasové sestavy svého týmu, musí zásadně vycházet z jejich objektivně individuálních aktuálních výkonů a projevů (např. psychická vyrovnanost, motivace do zápasu, schopnost okamžité reakce na změnu taktiky podle aktuálních pokynů trenéra, kterou trenér upravuje s vývojem zápasu atd.) a nikoliv podle jejich specializovaných postů při hře v obraně.

3. Seznam použitých zdrojů:

Cyhelský, Lubomír a Souček, Eduard. 2009. *Základy statistiky*. Praha : Vysoká škola finanční a správní, o.p.s. v edici EUPRESS, 2009. ISBN 978-807408-013-5.

Dovalil, Josef a kolektiv. 2012. *Výkon a trénink ve sportu*. Praha : Olympia Praha a. s., 2012. ISBN 978-80-7376-326-8.

Hindls, Richard, a další. 2007. *Statistika pro ekonomy*. Praha : Professional Publishing, 2007. ISBN 978-80-86946-43-6.

Liss, Vladimír a Východský, Pavel. 2014. Pravidla softballu 2014 - 2017. [Online] Česká softballová asociace, 2014. [Citace: 10. únor 2015.] http://www.softball.cz/download/2014/Pravidla_2014-2017_final_20140228.pdf.

Novovičová, Jana. 2006. *Pravděpodobnost a matematická statistika*. Praha : Vydavatelství ČVUT, 2006. ISBN 80-01-01980-2.

Süss, Vladimír. 2003. *Softball a Baseball*. Praha 7 - Holešovice : Grada Publishing, 2003. ISBN 80-247-0658-X.