

**Univerzita Pardubice
Fakulta ekonomicko-správní
Ústav matematiky**

**Využití logistické regrese a Gompertzovy křivky
v ekonomické praxi**

Bc. Ondřej Hurt

**Diplomová práce
2012**

ZADÁNÍ DIPLOMOVÉ PRÁCE

(PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

Jméno a příjmení: **Bc. Ondřej Hurt**
Osobní číslo: **E10171**
Studijní program: **N6208 Ekonomika a management**
Studijní obor: **Ekonomika a management podniku**
Název tématu: **Využití logistické regrese a Gompertzovy křivky
v ekonomické praxi**
Zadávací katedra: **Ústav matematiky**

Z á s a d y p r o v y p r a c o v á n í :

Cílem práce je na získaných datech aplikovat růstový model a predikovat budoucí hodnoty na základě získané funkce.

Diplomová práce bude obsahovat:

- vymezení růstových modelů
- definice logistické a Gompertzovy regrese
- význam parametrů regresní funkce
- aplikace růstových modelů v praxi.

Rozsah grafických prací: –
Rozsah pracovní zprávy: cca 30 stran
Forma zpracování diplomové práce: tištěná/elektronická

Seznam odborné literatury:

MARSHALL, Albert W. Life Distributions: Structure of nonparametric, semiparametric and parametric families. New York : Springer Science. 2007. 782 s. ISBN 978-0-387-20333-1

MATYS, Jaroslav. Gompertzova rozvojová křivka. Praha : Institut řízení, 1983. 88 s.

NUNES, Paul. Jumping the S-curve: How to beat the growth cycle, get on top, and stay there. Boston : Harvard Business School Publishing. 2011. 270 s. ISBN 978-1-4221-7558-3.

SEBER, George A. F. Nonlinear regression. New Jersey : John Wiley and Sons, Inc. 2003. 792 s. ISBN 978-0-471-47135-6

VYSEKALOVÁ, Jitka. Chování zákazníka: jak odkrýt tajemství "černé skříňky". 1. vyd. Praha : Grada, 2011. 356 s. ISBN 978-80-247-3528-3.

Vedoucí diplomové práce: doc. PaedDr. Jana Kubanová, CSc.
Ústav matematiky

Datum zadání diplomové práce: 30. června 2011

Termín odevzdání diplomové práce: 30. dubna 2012

doc. Ing. Renáta Myšková, Ph.D.

děkanka

L.S.

doc. Ing. Marcela Kožená, Ph.D.

vedoucí ústavu

V Pardubicích dne 4. srpna 2011

PROHLÁŠENÍ

Prohlašuji, že jsem tuto práci vypracoval samostatně. Veškeré literární prameny a informace, které jsem v práci využil, jsou uvedeny v seznamu použité literatury.

Byl jsem seznámen s tím, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorský zákon, zejména se skutečností, že Univerzita Pardubice má právo na uzavření licenční smlouvy o užití této práce jako Školního díla podle § 60 odst. 1 autorského zákona, a s tím, že pokud dojde k užití této práce mnou nebo bude poskytnuta licence o užití jinému subjektu, je Univerzita Pardubice oprávněna ode mne požadovat přiměřený příspěvek na úhradu nákladů, které na vytvoření díla vynaložila, a to podle okolností až do jejich skutečné výše.

Souhlasím s prezenčním zpřístupněním své práce v Univerzitní knihovně Univerzity Pardubice.

V Pardubicích dne 11. 3. 2012

Bc. Ondřej Hurt

PODĚKOVÁNÍ:

Tímto bych rád poděkoval paní doc. PaedDr. Janě Kubanové, CSc. za vedení diplomové práce, odbornou pomoc, připomínky při jejím vypracování a poskytnuté materiály, které byly využity při zpracování diplomové práce. Dále bych chtěl poděkovat svým rodičům za podporu, pomoc a trpělivost, kterou mi poskytli po dobu celého studia.

ANOTACE

Práce je zaměřena na aplikaci růstových modelů na ekonomické časové řady. Zabývá se použitím logistické a Gompertzovy regrese, rozбором jejich parametrů, výběrem nejvhodnější metody odhadu parametrů a následné predikce budoucích hodnot. Pro účely práce jsou využity programy MS Excel a Statistica 10.

KLÍČOVÁ SLOVA

Gompertzova křivka, logistická regrese, nelineární modely, odhad parametrů, růstové modely

TITLE

Usage of logistic regression and the Gompertz curve in economic practice

ANNOTATION

The work is focused on application of growth models for economic timeseries. It includes the usage of logistic regression and the Gompertz regression, analysis of their parameters, selecting the most appropriate methods of parameter estimation and prediction of future values. MS Excel and Statistica 10 are used for purposes of this work.

KEYWORDS

Gompertz curve, logistic regression, nonlinear models, parameter estimation, growth models

OBSAH

ÚVOD	10
1 NELINEÁRNÍ REGRESE	12
1.1 VYMEZENÍ GOMPERTZOVY KŘIVKY	13
1.1.1 <i>Vlastnosti Gompertzovy křivky</i>	13
1.1.2 <i>Parametry Gompertzovy křivky</i>	14
1.1.3 <i>Historie Gompertzovy regrese</i>	15
1.2 VYMEZENÍ LOGISTICKÉ KŘIVKY	16
1.2.1 <i>Vlastnosti logistické křivky</i>	16
1.2.2 <i>Parametry logistické křivky</i>	17
2 PŘEDPOKLADY POUŽITÍ RŮSTOVÉHO MODELU	19
3 POSTUP PŘI APLIKACI MODELU	21
3.1 ANALÝZA DAT	21
3.2 ODHAD PARAMETRŮ A SESTROJENÍ PREDIKČNÍ FUNKCE	21
3.3 VOLBA VHODNÉHO MODELU	22
3.3.1 <i>Interpolační kritéria</i>	22
3.3.2 <i>Extrapolací kritéria</i>	24
3.4 PROGNOZA HODNOT	24
4 METODY ODHADOVÁNÍ PARAMETRŮ	25
4.1 METODA ČÁSTEČNÝCH SOUČTŮ	25
4.2 METODA DÍLČÍCH PRŮMĚRŮ	26
4.3 METODA VYBRANÝCH BODŮ	26
4.4 ITERAČNÍ METODY	26
5 PODROBNĚJŠÍ ROZBOR MODELŮ Z EKONOMICKÉHO HLEDISKA	28
5.1 ŽIVOTNÍ CYKLUS VÝROBKU	28
5.2 EKONOMICKÝ A SOCIÁLNÍ VÝZNAM PARAMETRŮ	31
5.2.1 <i>Parametry u Gompertzovy regrese</i>	32
5.2.2 <i>Parametry u logistické regrese</i>	32
6 PRAKTICKÁ ČÁST	34
6.1 ILUSTRATIVNÍ PŘÍKLAD Č. 1	34
6.1.1 <i>Odhad parametrů Gompertzovy regrese</i>	34
6.1.2 <i>Predikce budoucích hodnot pomocí Gompertzovy rovnice</i>	47
6.1.3 <i>Extrapolace dat u Gompertzovy regrese</i>	47
6.2 ILUSTRATIVNÍ PŘÍKLAD Č. 2	49
6.2.1 <i>Odhad parametrů logistické regrese</i>	49
6.2.2 <i>Predikce budoucích hodnot pomocí logistické rovnice</i>	56
6.2.3 <i>Extrapolace dat u logistické regrese</i>	56
6.3 POROVNÁNÍ RŮSTOVÉHO A JEDNODUCHÉHO LINEÁRNÍHO MODELU	57
7 VYHODNOCENÍ JEDNOTLIVÝCH METOD	62
7.1 PŘEHLED VÝSLEDKŮ GOMPERTZOVY REGRESE	62
7.2 PŘEHLED VÝSLEDKŮ LOGISTICKÉ REGRESE	63
8 ZÁVĚR	65
POUŽITÁ LITERATURA	66

SEZNAM TABULEK

Tabulka 1: Ilustrativní příklad č. 1 - tabulka hodnot	35
Tabulka 2: Interpolační kritéria č. 1 - ilustrativní příklad č. 1	38
Tabulka 3: Optimalizace Gompertzovy regrese	40
Tabulka 4: Statistica 10 - výstup č. 1	43
Tabulka 5: Statistica 10 - výstup č. 2	44
Tabulka 6: Interpolační kritéria č. 2 - ilustrativní příklad č. 1	46
Tabulka 7: Extrapolace dat - náhodný výběr č. 1	47
Tabulka 8: Ilustrativní příklad č. 2 - tabulka hodnot	50
Tabulka 9: Interpolační kritéria č. 3 - ilustrativní příklad č. 2	52
Tabulka 10: Statistica 10 - výstup č. 3	54
Tabulka 11: Interpolační kritéria č. 4 - ilustrativní příklad č. 2	55
Tabulka 12: Extrapolace dat - náhodný výběr č. 2	56
Tabulka 13: Vyhodnocení Gompertzovy regrese	62
Tabulka 14: Vyhodnocení logistické regrese	63

SEZNAM OBRÁZKŮ

Obrázek 1: Životní cyklus výrobku	29
Obrázek 2: Nastavení řešitele	41
Obrázek 3: Zápis Gompertzovy rovnice	43
Obrázek 4: Zápis logistické regrese	53

SEZNAM GRAFŮ

Graf 1: Gompertzova křivka	15
Graf 2: Logistická křivka	18
Graf 3: Prodeje hry CoD: MW 3	30
Graf 4: Čtvrtletní prodeje PS 3	31
Graf 5: Čtvrtletní prodeje NDS	34
Graf 6: Porovnání metod odhadů parametrů u GR	37
Graf 7: GR pomocí řešitele	42
Graf 8: Statistica 10 - výstup č. 3	44
Graf 9: Počet prodaných mobilních telefonů v Rusku	49
Graf 10: Porovnání metod odhadů parametrů u logistické regrese	51
Graf 11: Logistická regrese pomocí řešitele	53
Graf 12: Statistica 10 - výstup č. 5	54
Graf 13: Jednoduchý lineární model	58
Graf 14: Simulace u jednoduché lineární regrese	60

SEZNAM ZKRATEK

COD: MW 3	Call of Duty: Modern Warfare 3
Dol. Sp. Mez	Dolní spolehlivostní mez
GR	Gompertzova regrese
Hor. Sp. Mez	Horní spolehlivostní mez
LR	Logistická regrese
M.A.E.	Mean Absolute Error
M.A.P.E.	Mean Absolute Percentage Error
M.E.	Mean Error
M.P.E.	Mean Percentage Error
M.S.E.	Mean Square Error
MČS	Metoda částečných součtů
MDP	Metoda dílčích průměrů
MNČ	Metoda nejmenších čtverců
MVB	Metoda vybraných bodů
NDS	Nintendo DS
PS 3	Play Station 3
ŘEŠ	Řešitel
SSE	Reziduální součet čtverců
SSR	Vysvětlený součet čtverců
SST	Celkový součet čtverců
STA	Program Statistica 10

Úvod

Tématem diplomové práce je využití logistické regrese a Gompertzovy křivky v ekonomické praxi. Tyto regrese slouží k modelování časových řad, pro které je typický růstový trend. Původně byly tyto modely vyvinuty pro modelování vývoje lidské populace a živočišných druhů, ovšem vzhledem k vlastnostem a charakteristice daných modelů nacházejí uplatnění i v oblasti ekonomických věd.

V první části práce jsou definovány a podrobněji vymezeny výše uvedené růstové modely, vlastnosti jejich parametrů a další matematické charakteristiky. Dále jsou uvedeny veškeré předpoklady, které musí zkoumaný soubor splňovat, aby na něj mohl být aplikován některý z růstových modelů. Tato práce je zaměřena na ekonomické časové řady, proto se při určování jednotlivých předpokladů vychází z této skutečnosti. V případě, že bychom zkoumali časovou řadu, která je tvořena náhodným výběrem pocházejícím z jiné oblasti než ekonomické, v takovém případě by se předpoklady pro použití růstového modelu lišily. Nejčastějším příkladem jsou objemy prodeje určitého výrobku za určité časové úseky.

Dále je určen přesný postup, který je vhodný při zkoumání a následném modelování určitého příkladu. Prvním krokem tohoto postupu je odhadnutí parametrů vybrané funkce za použití dostupných metod. Druhým krokem je vyhodnocení na základě statistických kritérií a v případě určení vhodnosti modelu bývá posledním krokem odhad budoucích stavů sledované veličiny.

Samotný odhad parametrů lze provést výpočtem pomocí tří metod, které jsou podrobněji popsány v další části textu. Druhou možností je výpočet pomocí iteračních metod, jejich ruční výpočet je ovšem složitý, proto bude pro účely této práce využit statistický program Statistica 10, který při výpočtu parametrů nelineární regrese využívá právě tyto iterační metody.

Významy parametrů růstových modelů jsou ze statistického hlediska jasně určeny a v mnohých textech podrobně popsány. Ovšem u těchto parametrů můžeme hledat i ekonomický popřípadě sociální význam, kdy například u jednoho z parametrů, který ovlivňuje strmost křivky, můžeme hledat vlivy, které existují na daném trhu a mohou být obsaženy právě v daném parametru.

Praktická část obsahuje dva ilustrativní příklady, kde je využita jak aplikace logistické regrese, tak i aplikace Gompertzovy regrese. Veškerý výpočet se řídí předpoklady a postupy vymezenými v teoretické části.

Cílem této práce je podrobně popsat a definovat logistickou a Gompertzovu regresi, vytvořit tak ucelenou práci, která se podrobněji zabývá teoretickými základy nelineárních modelů a vymezit postup a nástroje samotného modelování vybraných příkladů. Dále si tato práce klade za cíl na vybraných příkladech provést výpočet několika dostupnými a uvedenými metodami, určit pomocí statistických kritérií nejvhodnější model a na jeho základě predikovat budoucí stav sledovaného jevu.

K dosažení těchto cílů je v práci několikrát využit výpočet nebo jiná úprava v programu MS Excel a také výpočet v programu Statistica 10. Jako výstup pro hodnocení modelu jsou využity grafy vytvořeny pomocí některého z výše uvedených programů. Interpretace výsledků je uvedena v přehledné tabulce v závěrečné kapitole.

1 NELINEÁRNÍ REGRESE

Modely je možné podle linearity rozčlenit takto:

lineární modely,

nelineární, linearizovatelné,

nelineární, nelinearizovatelné.

Nelineárním modelem rozumíme takový model, který neobsahuje lineární vztah mezi veličinou X a veličinou Y a neexistuje možnost tento model transformovat na lineární tvar. Pokud není vztah lineární, je závislost mezi vysvětlovanou a vysvětlující proměnnou popsána nelineární regresí. S tímto typem závislosti je možno se setkat ve fyzikálních, chemických nebo ekonomických aplikacích, přičemž poslední jmenované odvětví bude využito v této práci. Je důležitá znalost dat a rozbor dané situace, které přispějí ke zvolení vhodného modelu. Pokud jsou ovšem informace omezené, může se přistoupit ke klasickému lineárnímu modelu. Metoda použití lineárního modelu je podstatně jednodušší. [7]

Klasická situace řešení aplikace ať už lineárního nebo nelineárního modelu je taková, že jsou k dispozici naměřené nebo jinak získané hodnoty, zpravidla označované jako vysvětlovaná proměnná nebo náhodná veličina Y , a zároveň odpovídající vysvětlující proměnná, náhodná veličina X , v ekonomických ale i v jiných aplikacích bývá často na ose x zobrazena časová jednotka. Cílem je proložit soubor náhodné veličiny Y co nejvhodněji odpovídající funkcí. Pro sestavení takové funkce je potřeba mít k dispozici naměřené hodnoty z určitého výběru, a na jejich základě odhadnout parametry daného modelu. Postup při aplikaci daného modelu bude popsán v dalším textu.

V lineárních modelech, kde rovnici obecně zapisujeme ve tvaru $Y = \alpha + \beta f(x)$, je význam parametrů jasně vymezen. Parametr α určuje posunutí na ose y při hodnotě $x_0 = 0$. Parametr β je směrnici přímky, přičemž při záporné hodnotě β je přímka klesající, při kladné hodnotě β je přímka rostoucí a pokud $\beta = 0$, je přímka rovnoběžná s osou x . Význam parametrů v nelineárních modelech je odlišný. Parametry určují například rychlostní konstanty, materiálové charakteristiky, rychlosti růstu atd. [7]

V případě lineárních modelů se používá pro odhad parametrů α a β metoda nejmenších čtverců¹.

¹ Podrobný popis je uveden v knize KUBANOVÁ, J.: *Statistické metody pro ekonomickou a technickou praxi*, 3. vydání, strana 107

V případě nelineárních modelů nelze metodu nejmenších čtverců použít, proto se používá metod jiných. Mezi ně je možné zařadit například metodu dílčích průměrů nebo metodu částečných součtů, tyto a jiné budou uvedeny a vymezeny v dalším textu.

Mezi nelineární modely se řadí následující: [11]

Modifikovaná exponenciální funkce ve tvaru $Y = \alpha + \beta^x$

Modifikovaná mocninná funkce ve tvaru $Y = \alpha + x^\beta$

Modifikovaná posunutá exponenciální funkce ve tvaru $Y = \kappa + \alpha \beta^x$

Logistická funkce ve tvaru $Y = \frac{\kappa}{1 + \alpha\beta^x}$

Gompertzova funkce ve tvaru $Y = \kappa\alpha^{\beta^x}$

Předmětem této diplomové práce budou nelinearizovatelné modely a to konkrétně poslední dva uvedené modely, logistická funkce a Gompertzova funkce. Tyto modely jsou shodně označovány jako s-křivky díky specifickému tvaru své funkce. Dále bývají zařazovány do kategorie růstových popřípadě rozvojových modelů.

V následujících dvou kapitolách budou tyto dva modely definovány, budou určeny jejich parametry a průběhy funkcí.

1.1 Vymezení Gompertzovy křivky

Gompertzova regrese byla zavedena v roce 1865, kdy ji publikoval Angličan Benjamin Gompertz v knize *On the Nature of the Function Expressive of the Law of Human Mortality, and on a New Mode of Determining the Value of Life Contingencies*. Prvotně byla použita v pojišťovnictví při odhadu průběhu úmrtnosti osob v závislosti na věku. Vztah, který je základem této regrese, se označuje jako Gompertzův zákon.[11]

Gompertzova regrese se využívá nejen v demografii, ale i v ekonomii. Její využití v demografii spočívá v modelování a prognózování vývoje počtu obyvatel v čase. V této spojitosti bývá křivka označována jako rozvojová nebo růstová. Aplikace v ekonomii je spjata s produkcí nových výrobků či služeb, popřípadě s poptávkou po novém statku. Důležitou skutečností, která vybízí k použití této regrese, je prodejní životní cyklus statku, ať už výrobku nebo služby.

1.1.1 Vlastnosti Gompertzovy křivky

Gompertzova křivka není na rozdíl od logistické křivky symetrická. Větší množství hodnot Gompertzovy křivky se nachází za inflexním bodem. Inflexní bod je hranicí, kde se průběh

křivky mění z konvexního na konkávní. Jedním z parametrů křivky je parametr κ , který je horní mezí funkce, dále je možno se setkat s označením horní asymptota κ , saturační mez κ , nebo hranice nasycení. Je to hranice hodnot, které průběh funkce nepřekročí. [8]

Je několik možností, jak rovnici Gompertzovy funkce vyjádřit, v této práci se bude používat tento zápis:

$$Y = \kappa \alpha^{\beta^x}$$

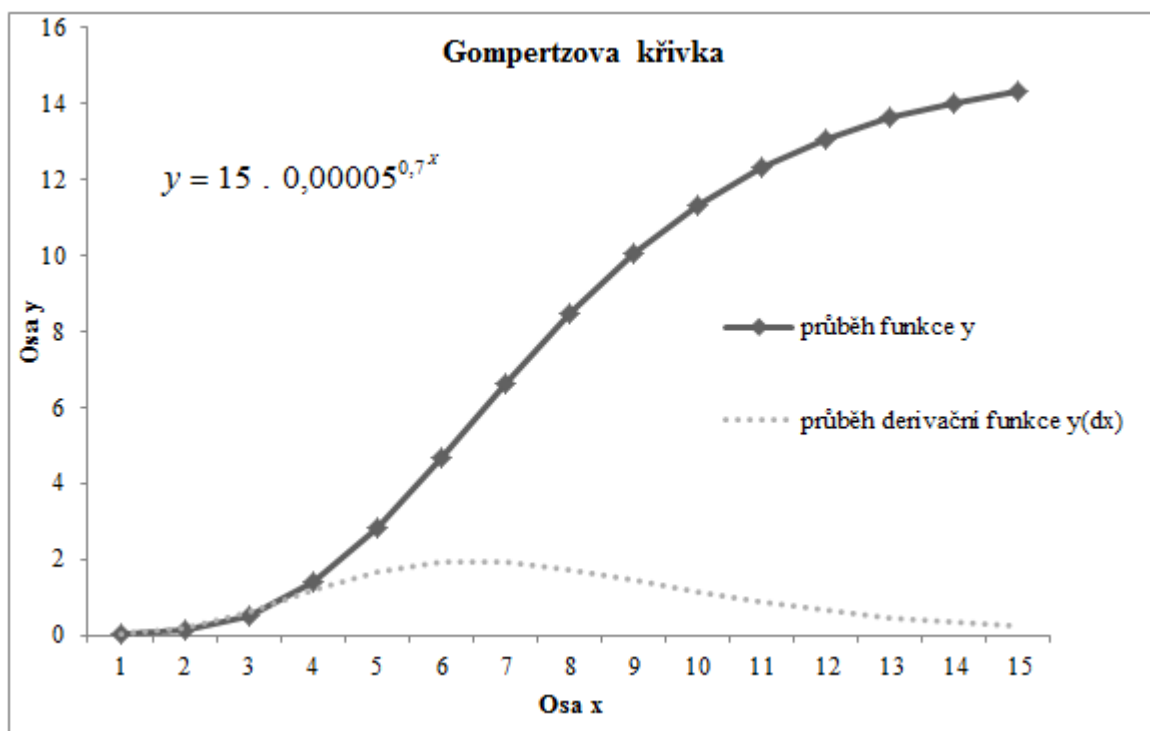
Rovnice má tři parametry, proto ji lze vymežit třemi hodnotami, nebo dvěma body a stavem nasycení. [14]

1.1.2 Parametry Gompertzovy křivky

Pro parametry této regresní funkce platí následující vztahy, $\kappa > 0$, $0 < \alpha < 1$ a $0 < \beta < 1$. Hodnota y vyjadřuje velikost sledovaného jevu, dále x je jednotka časového období a jak již bylo uvedeno, parametr κ udává saturační mez, tedy hranici nasycení. Parametr α vyjadřuje velikost sledovaného jevu v čase 0, tedy pro $x_0 = 0$, dosazením dostaneme vztah $y_0 = \kappa \alpha$, z čehož vyplývá, že $\alpha = \frac{y_0}{\kappa}$. Parametr β vyjadřuje míru strmosti, kdy platí takový vztah, že čím je parametr β menší, tím je průběh funkce strmější. [14], [10]

Průběh Gompertzovy funkce je charakteristický zvyšujícím se růstem okolo inflexního bodu, kdy jsou jednotlivé přírůstky největší. Průběh těchto diferencí zachycuje derivace dané rovnice. Diferencí rozumíme změnu y_i oproti y_{i-1} . Spojitou funkcí pro tyto rozdíly je derivace funkce. Derivaci také můžeme chápat jako rychlost změny. První derivace funkce je rovna výrazu $Y(dx) = \kappa \alpha^{\beta^x} \cdot \beta^x \ln \beta \cdot \ln \alpha$ a průběh této derivace je zobrazen na Grafu č. 1.

Na následujícím grafu je možno vidět průběh Gompertzovy křivky a její derivační průběh. Kdy je použita ilustrační funkce ve tvaru $y = 15 \cdot 0,00005^{0,7^x}$.



Graf 1: Gompertzova křivka

Zdroj: Vlastní zpracování v MS Excel

1.1.3 Historie Gompertzovy regrese

Zavedení Gompertzovy křivky je považováno za prvotní a zároveň nejúspěšnější matematický pokus o vyjádření závislosti úrovně úmrtnosti na věku. Následující vztah popisuje úmrtnost osob starších 20 let a je označován jako Gompertzův zákon.

$$\mu_x = R_0^{\alpha x}$$

Kde μ_x značí intenzitu úmrtnosti (rizikovou funkci) ve věku x , α a R jsou parametry tohoto vztahu. V pozdější době začal být Gompertzův zákon hojně užíván pro popis úmrtnosti laboratorních zvířat. Ve své práci Gompertz poznamenal, že kromě úmrtnosti exponenciálně narůstající s věkem může existovat i taková její složka, která je na věku nezávislá. „Je možné, že úmrtí může být důsledkem dvou obecně koexistujících příčin: jednotné riziko úmrtí bez předchozích predispozic a rostoucí riziko nebo také zvyšující se neschopnost odolat destrukci“ [6]. Pro pozdější analýzu úmrtnostních tabulek, které byly k dispozici, považoval Gompertz za možné se omezit pouze na exponenciální složku úmrtnosti. O 35 let později, v roce 1860, přidal další matematik, William Makeham, ke Gompertzově vztahu složku nezávislou na věku. Tento člen, obvykle značen písmenem A , bývá nazýván jako Makehamův parametr. Nově byl tak zformulován vztah nazýván jako Gompertz-Makehamův zákon v následujícím zápisu:

$$\mu_x = A + R_0^{ax}$$

Parametr A je na věku nezávislá komponenta úmrtnosti, kterou můžeme nazvat jako backgroundová složka úmrtnosti. Druhý člen R_0 je část úmrtnosti závislá na věku, která je dnes nazývána jako senescenční složka úmrtnosti. Z dosud uvedeného je patrné, že složka úmrtnosti závislá na věku má exponenciální průběh. Ve speciálním případě, kdy backgroundová složka může být zanedbána (např. v dobrých laboratorních podmínkách nebo v současné době v průmyslově vyspělých státech), roste celková intenzita úmrtnosti exponenciálně s věkem, tedy v souladu s Gompertzovým zákonem.[5]

1.2 Vymezení logistické křivky

Logistická křivka byla představena Raymondem Pearlem a Lowellem Reedem v roce 1920 jako vztah popisující růst lidí a zvířat. Je proto někdy označována jako „*Pearl-Reed curve*“. V následujících letech se na tento model snesla kritika a to zejména kvůli tvrzení R. Pearla, který uvedl, že tento model popisuje zákon růstu dané populace. S tímto tvrzením nesouhlasila celá řada biologů, ekonomů a statistiků. Nicméně se tento model objevil v 30. letech 19. století jako jeden z hlavních modelů experimentální biologie a ve svých různých modifikacích zůstává důležitou součástí moderních věd dodnes. [9]

Logistická křivka se stejně jako Gompertzova křivka řadí mezi růstové neboli rozvojové modely. Průběh této křivky vyjadřuje rozvoj určité populace v čase. Díky svému tvaru podobnému písmenu S se označuje společně s Gompertzovou křivkou jako s-křivka.

1.2.1 Vlastnosti logistické křivky

Pro křivku je charakteristický symetrický průběh trendu, což znamená, že jak na levé tak i na pravé straně od inflexního bodu se nachází stejný počet hodnot. Inflexní bod se nachází v bodě $x_i = \frac{-\ln \alpha}{\ln \beta}$ a stejně jako předchozí regrese má kladnou asymptotu κ . [8]

I zde je několik možností, jak zapsat tvar logistické funkce. V této práci se bude vycházet z následujícího zápisu:

$$Y = \frac{\kappa}{1 + \alpha\beta^x}$$

Tato rovnice má tři parametry, proto jsou možnosti jejího vymezení obdobné jako u Gompertzovy křivky.

Průběh logistické funkce je rovněž charakteristický zvyšujícím se růstem okolo inflexního bodu. Průběh těchto diferencí zachycuje derivace dané rovnice. První derivace funkce je

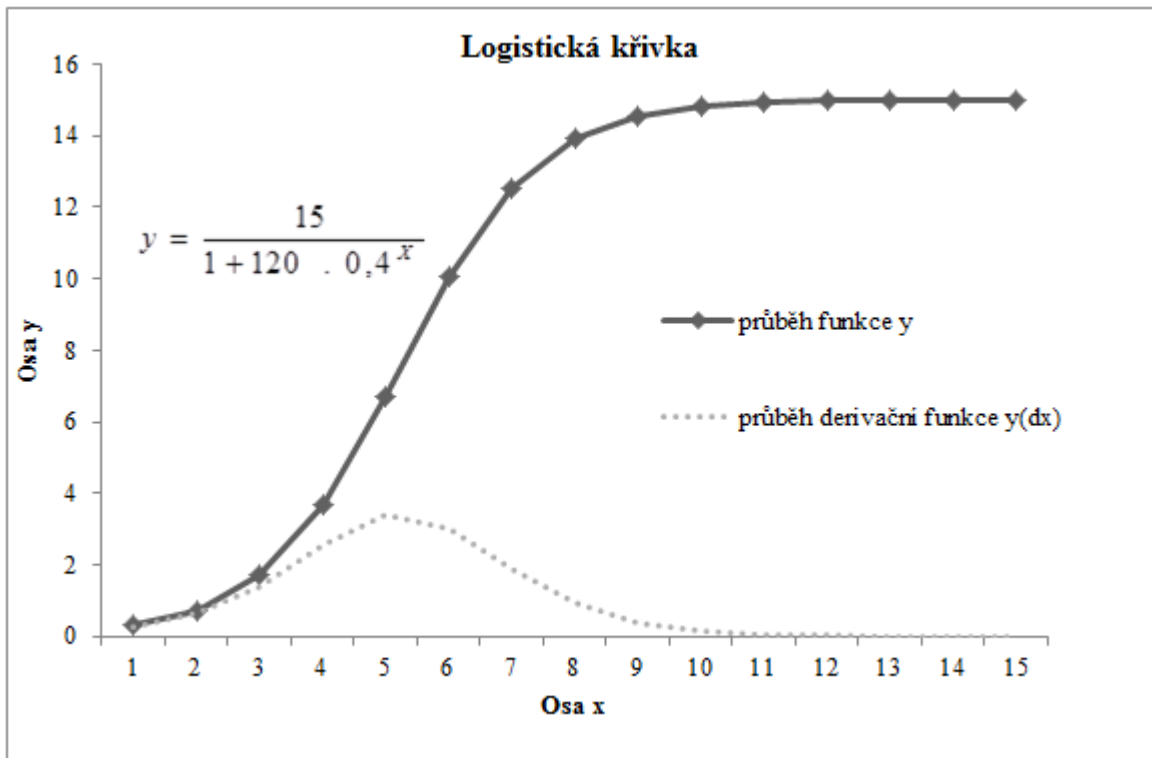
rovna výrazu $Y(dx) = \frac{-\kappa\alpha\beta^x \cdot \ln \beta}{(1 + \alpha\beta^x)^2}$. Derivační průběh je zobrazen na grafu č. 2. Tento průběh první derivace je pro s-křivky typický. Zpočátku se diference zvětšují pozvolně, a čím více se blíží k inflexnímu bodu, tak se hodnoty diferencí zvětšují. Po překročení inflexního bodu se hodnoty diferencí postupně zmenšují a čím blíže je rovnice křivky k asymptotě, tím více se hodnoty diferencí zmenšují a blíží se k nule. Průběh derivační funkce tedy vypovídá o specifickém tempu růstu dané nelineární regrese.

1.2.2 Parametry logistické křivky

Saturační mez κ může nabývat jakýchkoli kladných hodnot, koeficient α může nabývat hodnot větších než 1 a koeficient β se pohybuje v uzavřeném intervalu $\langle 0; 1 \rangle$.

Význam parametrů vzhledem k tvaru rovnice je odlišný. Vliv na strmost křivky má koeficient β a platí, že čím je hodnota koeficientu nižší, tím je křivka strmější. Hodnota funkce v čase $x_0 = 0$ je rovna výrazu $\frac{\kappa}{1 + \alpha}$. [46]

Na následujícím grafu je možno vidět průběh logistické křivky a její rychlost změny neboli derivační průběh. Kdy je použita ilustrační funkce ve tvaru $y = \frac{15}{1 + 120 \cdot 0,4^x}$.



Graf 2: Logistická křivka

Zdroj: Vlastní zpracování v MS Excel

2 PŘEDPOKLADY POUŽITÍ RŮSTOVÉHO MODELU

Pro použití jednoho z růstových modelů, a to ať už Gompertzova nebo logistického trendu, je potřeba, aby sledované hodnoty a charakter sledovaného jevu splňoval určité podmínky. Pokud by některé z následujících podmínek nebyly splněny, pozbývala by aplikace jednoho z modelů na naměřená data a následná predikce budoucích hodnot smyslu.

Podmínky pro použití růstového modelu: [14]

- Sledovaný jev (statek) je v době vstupu na trh nový a inovativní, je pozitivně odlišen od podobných statků
 - V případě výrobku je výrobek kvalitnější a funkce, které plní, plní lépe než jeho předchůdci či konkurenti.
 - Křivka sleduje časový úsek od uvedení na trh až po upevnění stabilní pozice na trhu.
- Statek po celou dobu sledování plní stejnou funkci
 - U výrobku mohou v průběhu jeho životnosti nastat různé inovace a změny, ovšem tyto změny nemají větší vliv na základní funkce, které výrobek plní.
- Je sledováno celkové množství výrobků, které vstupují do užívání
 - Výrobky, které nahrazují vyřazené objekty, se již nově nesledují a nevstupují jako měřené hodnoty.
- Vymezení určité geografické a odvětvové oblasti
 - Sleduje se prodej výrobků v určité geografické oblasti, například prodej aut v České republice.
 - Zároveň se vymezí i výrobky spadající do stejné kategorie, sleduje se například prodej aut určité značky a druhu. To znamená, že se nesledují prodeje aut od dvou různých výrobců.
- Absence umělých zásahů do výroby a odbytu
 - Výroba a následný prodej probíhaly ve sledovaném období přirozeným způsobem, znamená to například nepřítomnost většího snížení ceny nebo jiného marketingového zásahu.
 - Tato podmínka je velmi obsáhlá a komplikovaná a záleží na subjektivním posouzení každého jedince a zároveň znalosti sledovaného jevu. Je několik

vstupů (zásahů), které mohou větším či menším způsobem ovlivnit získávané hodnoty. Vstupy můžeme rozdělit do dvou skupin podle toho, jaký vliv na sledovaný jev mají, buď posílí prodej (růst křivky), nebo utlumí prodej (vytvoření pomyslné saturační meze). Mezi posilující vstupy můžeme například zařadit již zmiňovanou cenovou změnu, dále velký příliv obyvatel do sledovaného regionu (tito noví obyvatelé představují potencionální poptávku), nebo větší snížení počtu konkurentů, kteří prodávají stejný výrobek. Mezi utlumující vstupy můžeme zařadit inverzně příchod nových konkurentů do regionu, nebo nucené zvýšení prodejní ceny.

- Shrnutím této podmínky může být výsledný předpoklad, že by na daném vymezeném trhu, měly být splněny předpoklady pro dokonalou konkurenci a s tím související fungování trhu. (Zde patří předpoklady, jako jsou žádné bariéry vstupu do odvětví, dokonalé informace pro kupující, stejné podmínky pro výrobce atd.)

- Periodičnost

- Vhodným časovým úsekem, za který se hodnoty získávají, bývá v ekonomické aplikaci jeden rok, pokud by totiž bylo zvoleno období kratší, než je jeden rok, mohla by získané hodnoty ovlivnit sezónnost během roku. V takovém případě by bylo možné vidět už i při pouhé ilustraci pomocí grafu, že použití růstového modelu, není vhodné.

3 POSTUP PŘI APLIKACI MODELU

Při samotném řešení daného problému, kterým je aplikace vhodného modelu na získaná data, je možno vymežit následující postup:

1. Seznámení s daty,
2. odhad parametrů a sestavení predikční funkce,
3. volba vhodného trendu,
4. prognóza hodnot.

3.1 Analýza dat

Velmi důležitou částí v modelování určité situace je dobrá znalost problému a co nejlepší porozumění daným hodnotám. Jedná se o první krok při modelovacím procesu, který často rozhoduje o tom, zda výsledek bude kvalitní a použitelný nebo nikoliv. Předpokladem je rozbor všech možných vstupů, které mohou potencionálně ovlivnit sledovaný jev. Těmto vstupům se podrobněji věnovala druhá kapitola.

Pokud by například nebyla věnována dostatečná pozornost tomuto kroku, mohlo by dojít k přehlédnutí určitého faktoru, který ovlivňuje (posiluje nebo utlumuje) budoucí hodnoty, a výsledný vymodelovaný trend by byl chybně zkonstruován.

Cílem je se co nejvíce seznámit s danou situací a tím předejít možným chybám, ke kterým by mohlo během postupu dojít. Čím větší pozornost bude věnována této části, tím je větší pravděpodobnost, že výsledný model bude co nejvíce kopírovat skutečnou situaci.

3.2 Odhad parametrů a sestavení predikční funkce

Při odhadování parametrů nelze postupovat stejným způsobem jako u lineárních nebo linearizovatelných modelů, kdy v obou případech lze využít metodu nejmenších čtverců (MNČ). Podstatou MNČ je proložení naměřených dat takovou přímkou, která co nejtěsněji přiléhá k daným hodnotám. Výsledkem je odhad parametrů α a β . [10]

V situaci, kdy chceme odhadnout parametry nelinearizovatelného modelu, se nám nabízí dvě možnosti. První možností je použití iteračních metod. Iterační metody jsou výpočetně složité a náročné, proto se s nimi můžeme setkat například v celé řadě statistických programů. Ovšem „ruční“ výpočet těchto metod by byl velmi složitý. Jednou z těchto iteračních metod je například Levenberg-Marquardtova metoda. Podstatou této metody je také metodika nejmenších čtverců, ovšem v komplikovanější formě. [12]

Druhou možností je použít metody odhadu parametrů, které se používají u jiného nelineárního modelu, a to u modifikované posunuté exponenciály, jejíž tvar funkce je $Y = \kappa + a \beta^x$. Mezi tyto metody odhadu řadíme metodu částečných součtů, metodu dílčích průměrů a metodu vybraných bodů. Tyto metody spolu s metodami iteračními budou podrobněji definovány v dalším textu.

Výsledkem těchto metod jsou požadované odhady neznámých parametrů. Odhadem těchto parametrů tedy získáme predikční funkci. Po získání této rovnice se zpravidla určí teoretické hodnoty a zhodnotí se, nakolik je predikční funkce kvalitní.

3.3 Volba vhodného modelu

Dalším krokem je volba vhodného modelu. Vzhledem k tomu, že předmětem této práce jsou pouze dva růstové modely, budeme se zabývat výběrem z těchto dvou s-křivek. Nebudeme tedy brát v úvahu ostatní lineární či linearizovatelné modely.

Při volbě mezi Gompertzovou a logistickou křivkou je důležité si uvědomit jejich základní rozdíl. Gompertzova křivka obsahuje větší množství hodnot za inflexním bodem, což je možné chápat také jako skutečnost, že časové období, kdy se křivka přibližuje ke své asymptotě, je delší než u logistické křivky.

Další možností je zvolit mezi modely na základě grafického znázornění obou trendů. Odhadneme parametry obou rovnic a spolu se skutečnými hodnotami y_i si zobrazíme pomocí grafu jak Gompertzovu funkci, tak i predikční logistickou funkci. Přičemž bude vybrána právě ta rovnice, která kopíruje co nejpřesněji skutečné naměřené hodnoty. V tomto případě bude vhodnější regrese vybrána na základě interpolačních a extrapolacních kritérií, která budou podrobněji popsána v následující podkapitole.

Pokud budeme na řešený problém nahlížet jako na časovou řadu, je možno vymezit členění na extrapolaci či interpolaci. Extrapolací dat rozumíme zaměření se na sestavení takového modelu, který bude co nejpřesněji predikovat budoucí hodnoty. Interpolací dat rozumíme takový postup, u kterého se zaměřujeme na co nejpřesnější proložení naměřených hodnot křivkou, zaměřujeme se tedy na již získané, minulé hodnoty. [8]

3.3.1 Interpolační kritéria [8]

Podstatou těchto kritérií je porovnání součtu popřípadě průměru čtverců odchylek skutečných hodnot od teoretických hodnot. Jedná se tedy o ukazatele míry přilnavosti dané křivky ke skutečným hodnotám. Žádoucí výsledek je takový, aby součet nebo průměr těchto

odchylek byl co nejmenší. Čím je ukazatel menší, tím je model kvalitnější a vhodně kopíruje skutečnost. Mezi použité ukazatele můžeme zařadit následující:

M. E. Mean Error – střední chyba odhadu, kde vycházíme ze vztahu:

$$M.E. = \frac{\sum_{t=1}^n (y_t - \hat{y}_t)}{n}$$

kde y_t – naměřená hodnota y ; \hat{y}_t – teoretická hodnota \hat{y} ; n – počet měření. Tato míra je rovna nule pokud model přesně kopíruje měřená data, nebo pokud dojde k takové kompenzaci kladných a záporných rozdílů mezi y_t a \hat{y}_t , že výsledná hodnota čitatele bude nula.

M. S. E. Mean Square Error – střední čtvercová chyba odhadu:

$$M.S.E. = \frac{\sum_{t=1}^n (y_t - \hat{y}_t)^2}{n}$$

Tato metoda měření je nejpoužívanější, v tomto případě díky umocnění každého rozdílu y_t a \hat{y}_t získáme pouze kladné hodnoty. Tento postup je pro nás samozřejmě vhodnější.

M. A. E. Mean Absolute Error – střední absolutní chyba odhadu

$$M.A.E. = \frac{\sum_{t=1}^n |y_t - \hat{y}_t|}{n}$$

M.A.P.E. Mean Absolute Percentage Error – střední absolutní procentní chyba odhadu

$$M.A.P.E. = \sum_{t=1}^n \left(\frac{|y_t - \hat{y}_t|}{y_t} \right) \cdot \frac{100}{n}$$

M.P.E. Mean Percentage Error – střední procentní chyba odhadu

$$M.P.E. = \sum_{t=1}^n \left(\frac{y_t - \hat{y}_t}{y_t} \right) \cdot \frac{100}{n}$$

3.3.2 Extrapolační kritéria [8]

Způsob použití extrapolačních kritérií spočívá v simulaci. Postup je takový, že se vynechají poslední naměřené hodnoty a na zbylé hodnoty se aplikuje model. Pomocí takto získaného modelu predikujeme budoucí hodnoty, přičemž čím více budou tyto hodnoty odpovídat vynechaným hodnotám, tím bude zvolený model vhodnější. Jako příklad extrapolačního ukazatele si můžeme uvést koeficient nesouladu, též známý jako Theilův koeficient.

Theilův koeficient má následující tvar:

$$T_H^2 = \frac{\sum_{j=1}^D (y_{N+j} - \hat{P}_j)^2}{\sum_{j=1}^D y_{N+j}^2}$$

kde

N – délka časové řady použita pro odhad (po zkrácení),

D – zkrácení časové řady,

\hat{P}_j - extrapolace na j období.

Pro výsledek extrapolace je žádoucí, aby hodnota Theilova koeficientu byla co nejbližší hodnotě 0, v tomto případě by predikované hodnoty odpovídaly hodnotám vynechaným.

3.4 Prognóza hodnot

Posledním krokem je prognóza budoucích hodnot. Právě poslední krok je stěžejní při ekonomických aplikacích. Například v případě modelování objemů prodeje určitého výrobku je předpověď budoucí poptávky zákazníků velice důležitá při řízení a rozhodování firmy. Pokud zná firma odhad budoucí poptávky, může tomuto odhadu přizpůsobit výrobu a veškeré další procesy a činnosti s tím související. Čímž může dojít k výraznému zefektivnění daných postupů a minimalizaci nákladů. Toto pravděpodobně bude hrát klíčovou roli v účasti dané firmy na trhu. Pokud je firma schopna odhadnout budoucí prodeje, dostává se do konkurenční výhody oproti svým obchodním protivníkům. Tento příklad je jen dalším důkazem důležitosti využívání statistických metod v ekonomickém odvětví.

4 METODY ODHADOVÁNÍ PARAMETRŮ

Jak již bylo napsáno dříve, k odhadování parametrů obou růstových modelů se může využít metod používaných u modifikované posunuté exponenciály $Y = \kappa + \alpha \beta^x$. Nejdříve je ovšem nutné jak logistickou, tak i Gompertzovu rovnici transformovat na tvar odpovídající tvaru modifikované posunuté exponenciály.

Transformace logistické funkce:

$$Y = \frac{\kappa}{1 + \alpha \beta^x}$$

Transformaci provedeme převrácením výrazu a obdržíme:

$$\frac{1}{Y} = \frac{1}{\kappa} + \frac{\alpha}{\kappa} \beta^x$$

Nyní je již rovnice upravena na požadovaný tvar.

Transformace Gompertzovy funkce:

$$Y = \kappa \alpha^{\beta^x}$$

Transformaci provedeme zlogaritmováním celé rovnice a obdržíme:

$$\ln Y = \ln \kappa + \beta^x \ln \alpha$$

Nyní je i Gompertzova rovnice upravena na požadovaný tvar. [11]

Dále budou popsány tři metody pro odhadování parametrů dané funkce. Kdy odhad parametru α , provedeme pomocí a , parametr β bude odhadnutý pomocí b , a parametr κ bude odhadnutý pomocí k .

4.1 Metoda částečných součtů [11]

Tato metoda spočívá v rozdělení sledovaného souboru o n pozorováních do 3 stejně početných částí o délce m . Pro dané 3 části určíme jejich součty, které označíme postupně S_1 , S_2 a S_3 . Pokud počet souhrnného pozorování neumožní rozdělení n pozorování do 3 vyvážených částí, zpravidla se vynechávají nejstarší pozorování. Odhady parametrů jsou určeny podle následujících vztahů:

$$b = \sqrt[m]{\frac{S_3 - S_2}{S_2 - S_1}}; \quad a = \frac{(b-1)(S_2 - S_1)}{b(b^m - 1)^2}; \quad k = \frac{1}{m} \left(S_1 - ab \frac{b^m - 1}{b-1} \right)$$

4.2 Metoda dílčích průměrů [11]

I zde je celý sledovaný soubor rozdělen do 3 stejných částí. U počátečních hodnot pozorování určíme průměr, který označíme jako S_D , kde D značí dolní hodnoty pozorování. U středních hodnot postupujeme stejně a průměr označíme S_P , kde P značí prostřední hodnoty pozorování. Nakonec tento průměr uděláme i u posledních známých hodnot pozorování a průměr označíme S_H , kde H značí horní hodnoty pozorování. Pomocí těchto průměrů spočítáme hodnotu koeficientu b podle vztahu:

$$b = \frac{n-5}{2} \sqrt{\frac{S_H - S_P}{S_P - S_D}}$$

Dále postupujeme jako v případě odhadu parametrů metodou nejmenších čtverců.

4.3 Metoda vybraných bodů [11]

Způsob této metody spočívá v určení tří hodnot z naměřených dat, které vybereme podle vztahů y_x , y_{x+m} a y_{x+2m} , kde zpravidla $x = 1$ a $m =$ přirozené číslo. Získané hodnoty dosadíme do rovnic, kdy obdržíme soustavu tří rovnic, jejímž řešením jsou odhady parametrů b , a , k .

$$b = \sqrt[m]{\frac{y_{x+2m} - y_{x+m}}{y_{x+m} - y_x}}; \quad a = \frac{y_{x+m} - y_x}{b^x (b^m - 1)}; \quad k = y_x - ab^x$$

4.4 Iterační metody

Levenberg-Marquardtův algoritmus je iterační technika, která vyhledá minimum funkce, jež je vyjádřena jako součet čtverců nelineárních funkcí. Tento algoritmus se stal standardní používanou technikou pro výpočet parametrů nelineárních funkcí. [47]

Využívá explicitní výpočet Hessovy matice, kdy se jedná o kombinaci Gauss-Newtonovy metody s metodou největšího spádu. Cílem této kombinace je vyvarovat se problémům způsobených Gauss-Newtonovou metodou. Pro Gauss-Newtonovu metodu platí následující vztah: $A \delta^{\rightarrow} x = -\nabla f|_{\vec{x}^{(k)}}$.

V Levenberg-Marquardtově metodě nahradíme matici A maticí A' .

$$A' = \begin{cases} a'_{ij} = a_{ij}(1 + \lambda) \\ a'_{ij} = a_{ij} \end{cases}; i \neq j$$

Pokud $\lambda \rightarrow 0$ dostáváme Gauss-Newtonovu metodu a pro $\lambda \rightarrow \infty$ dostáváme malý krok ve směru spádu. Postup vypadá takto:

1. Stanovíme vektor $\vec{x}^{(0)}$ a hodnotu $f(\vec{x}^{(0)})$.
2. Udáme hodnotu λ , například $\lambda = 0,001$, pro $k = 0$.
3. Použijeme vztah $A' \Delta \vec{x} = -\nabla f \big|_{\vec{x}^{(k)}}$.
4. Vypočteme hodnotu $\vec{x}_p = \vec{x}^{(k)} + \Delta \vec{x}$ a $f(\vec{x}_p)$.
5. Pokud $f(\vec{x}_p) \geq f(\vec{x}^{(k)})$, krok zamítneme $\rightarrow \lambda := 10\lambda$.
6. Pokud $f(\vec{x}_p) < f(\vec{x}^{(k)})$, krok přijmeme $\vec{x}^{(k+1)} = \vec{x}_p$ a $\lambda := 0,1\lambda$ a $k := k + 1$. Dále cyklus opakujeme od kroku č. 3.

Levenberg-Marquardtova metoda se často využívá při odhadování parametrů nelineární regrese. V takových případech minimalizujeme funkci $X^2 = \sum_{i=1}^N [y_i - y(x_i, \vec{a})]^2 / \sigma_i^2$ vzhledem k \vec{a} .

Pak $\frac{\partial^2 X^2}{\partial a_k \partial a_l} = 2 \sum_{i=1}^N \frac{1}{\sigma_i^2} \left[\frac{\partial y(x_i, \vec{a})}{\partial a_k} \frac{\partial y(x_i, \vec{a})}{\partial a_l} - (y_i - y(x_i, \vec{a})) \frac{\partial^2 y}{\partial a_k \partial a_l} \right]$. Zde je možné část sumy,

kde se vyskytují druhé parciální derivace y , zanedbat, jelikož pro řešení stačí vypočítat matici A přibližně. [12] Levenberg-Marquardtova metoda je tedy jakousi nadstavbou Gauss-Newtonovy metody.

Jak již bylo napsáno výše, tato metoda je výpočetně komplikovaná, proto bude v praktické části použit výpočetní software Statistica 10, který tyto dva algoritmy využívá.

5 PODROBNĚJŠÍ ROZBOR MODELŮ Z EKONOMICKÉHO HLEDISKA

V této kapitole bude podrobněji ukázána podobnost mezi životním cyklem výrobku a růstovými modely a poté budou vymezeny parametry jednotlivých regresí z ekonomického hlediska.

5.1 Životní cyklus výrobku

Jak již bylo uvedeno, růstové modely nacházejí vyžití v oblasti ekonomie zejména při odhadování budoucích prodejů nových výrobků, a to především díky svému specifickému tvaru. Tento tvar jak u logistické tak i u Gompertzovy regrese odpovídá počátečním fázím životního cyklu nového výrobku.

Fáze životního cyklu jsou následující [2]:

- Uvedení

Výrobek je uveden na trh, počáteční prodeje jsou velmi nízké a náklady spojené s tímto produktem často převyšují příjmy plynoucí z prodejů

- Růst

Výrobek si našel své místo na trhu a stává se oblíbeným mezi zákazníky, prodeje se každým dalším obdobím navyšují a příjmy zpravidla přesahují v této fázi náklady

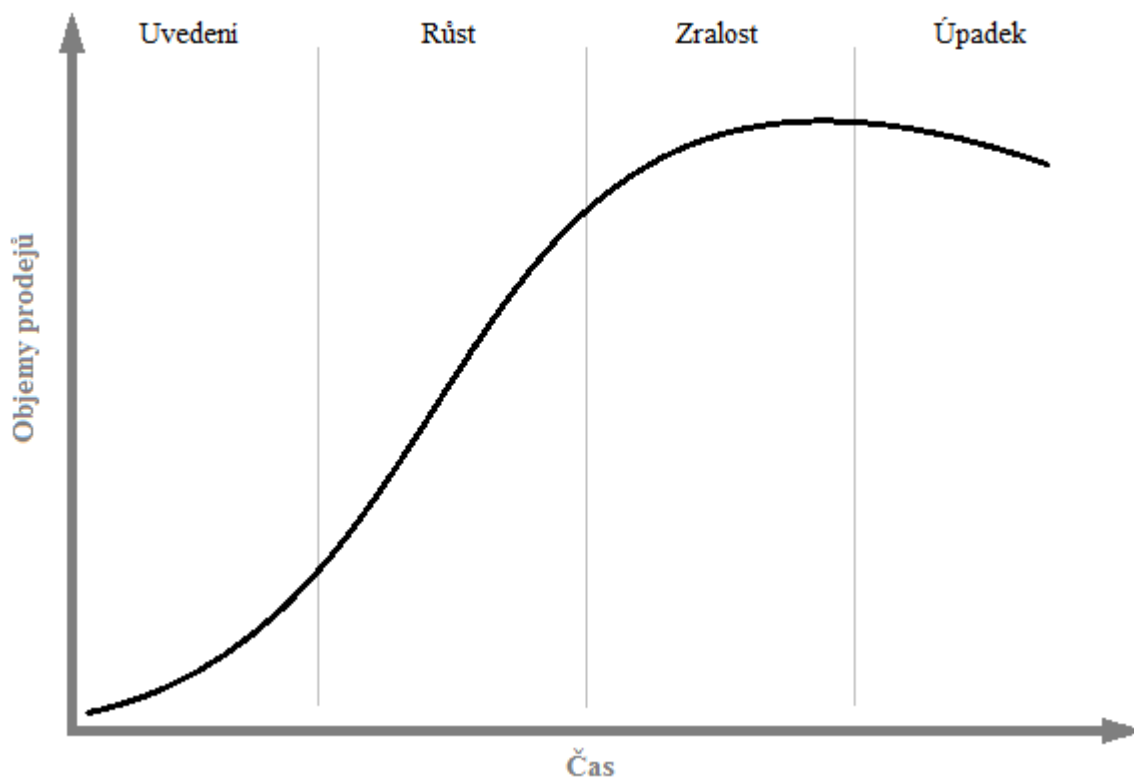
- Zralost

Objemy prodejů dosahují svého maxima, kdy výrobek je již dobře známý a oblíbený mezi zákazníky

- Úpadek

Faktor, který je rozhodující v této poslední fázi, je čas. Výrobek je zastaralý a na trhu bývá zpravidla nahrazen novějšími produkty. Prodeje klesají a pro prodejce nastává čas k zaměření se na jiný produkt.

Na následujícím obrázku můžeme vidět grafické znázornění jednotlivých fází životního cyklu výrobku.



Obrázek 1: Životní cyklus výrobku

Zdroj: vlastní zpracování v MS Excel, [4]

Jak napovídá obrázek č. 1, fáze uvedení, růstu a zralosti odpovídají tvaru růstových modelů. Nejdůležitějším faktorem v jednotlivých fázích je narůstající obliba a reference mezi zákazníky. Jak se výrobek stává populárnější mezi lidmi, rostou i objemy prodeje.

Oblíbenost produktu však není nejdůležitějším faktorem. Existují i produkty, které splňují předpoklady pro aplikaci růstových modelů, ovšem drobné rozdíly například ve vlastnostech výrobku nebo ve způsobech jeho propagace zcela mění jeho fáze životního cyklu. Na tomto základě můžeme výrobky rozdělit do tří skupin:

- Inovativní,
- očekávané,
- sezónní.

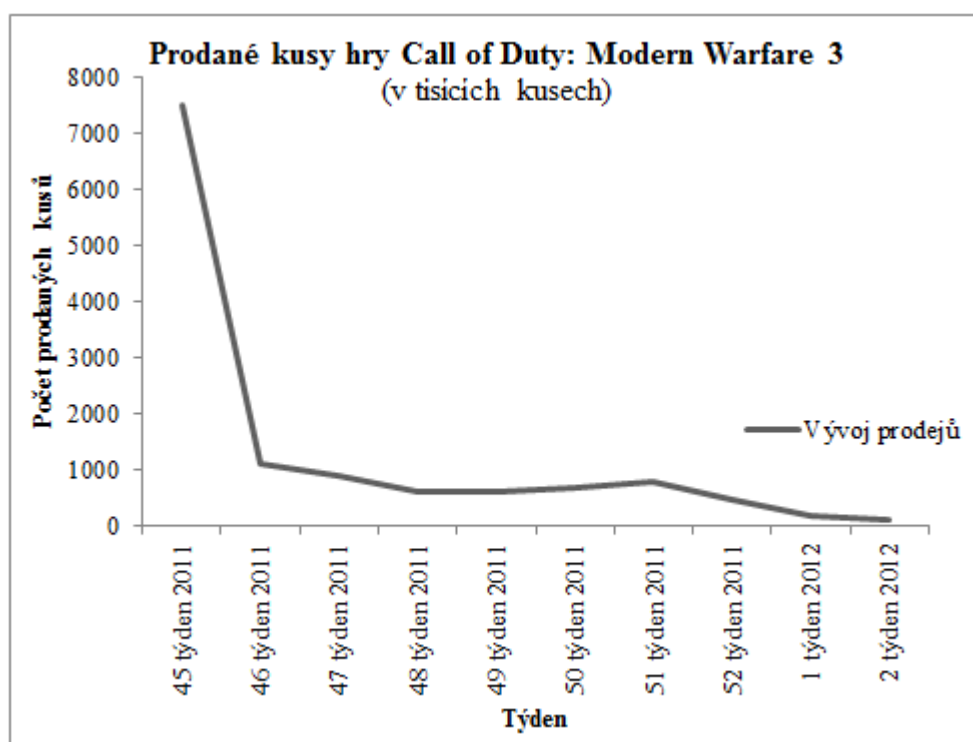
Inovativní výrobky

Do této skupiny patří výrobek zcela neznámý, který je ve své době inovativní a přináší zákazníkovi něco nového. Zákazník není obeznámen s přístupem nového výrobku na trh, nebo existuje určitá forma propagace, která ovšem zákazníkovi zpravidla nenapoví, jaký výrobek ve skutečnosti bude. Pro tyto výrobky je odpovídající životní cyklus zobrazen na

obrázku č. 1. Tento výrobek a jeho vystupování na trhu splňuje veškeré předpoklady vymezené v kapitole 2. Jako příklady můžeme uvést celosvětové prodeje herní konzole Nintendo DS, mobilních telefonů, stolních počítačů a mobilních telefonů iPhone.

Očekávané výrobky

Tento výrobek má zpravidla své předchůdce, proto zákazník očekává, jaký užitek mu případná koupě výrobku přinese. Výrobek bývá inovativní a nese sebou něco nového, ovšem není neznámý. Prodeje tohoto výrobku zpočátku prudce rostou a s postupem času klesají. Dobrým příkladem mohou být počítačové hry, kde můžeme jako příklad a ilustraci na grafu č. 3 uvést prodeje počítačové hry Call of Duty: Modern Warfare 3.



Graf 3: Prodeje hry CoD: MW 3

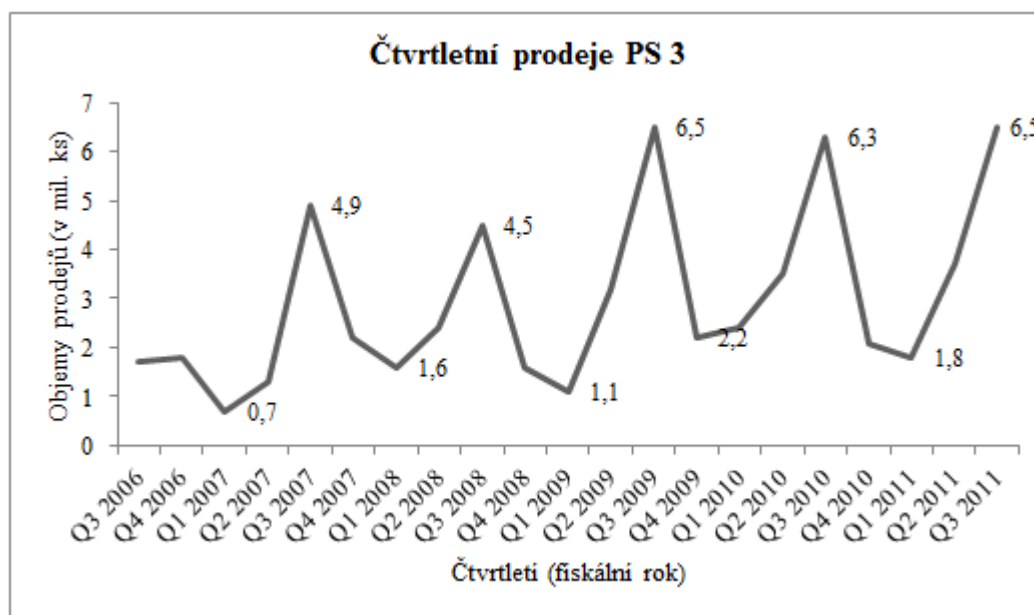
Zdroj: vlastní zpracování v MS Excel, [3]

Na předchozím grafu můžeme sledovat, že při vydání hry ve 45 týdnu roku 2011 došlo k největším prodejem a v dalších týdnech byly prodeje výrazně nižší.

U tohoto druhu produktu je zajímavé sledovat vývoj počtu uskutečněných předobjednávek, kdy má trend obdobný průběh jako u růstových modelů, ovšem vzhledem ke své podstatě, která je u předobjednávek odlišná od klasických výrobků, není možné aplikovat růstový model na takovýto druh produktu. Hlavní rozdíl spočívá v jasně určené a omezené době, po kterou se produkt nabízí. Jakmile dojde k vydání například dané počítačové hry, prodej předobjednávek je ukončen.

Sezónní výrobky

Tento výrobek se od ostatních liší pouze v odlišné propagaci, která je zapříčiněna zpravidla vyšší cenou produktu. Sezónnost prodeje spočívá v zaměření prodejce na propagaci na období 4. čtvrtletí kalendářního roku. Výrobek je také nový a inovativní, ovšem vyšší cena, nebo užitek, který pro zákazníka není nezbytný, zapříčiní sezónnost v prodeji. Typickým příkladem je zboží jako spotřební elektronika, kam můžeme zařadit mobilní telefony vyšší třídy, či herní konzole. Pro ilustraci je na grafu č. 4 zobrazen vývoj prodejů herní konzole Sony Playstation 3.



Graf 4: Čtvrtletní prodeje PS 3

Zdroj: vlastní zpracování v MS Excel, [45]

Na grafu č. 4 můžeme sledovat každoroční nárůst prodejů vždy ve 3. čtvrtletí fiskálního roku, tedy v období Vánoc. Prodeje v těchto čtvrtletích dosahovaly až 7 milionů prodaných kusů od svého uvedení na trh v roce 2006. Herní konzole Sony Playstation 3 (PS 3) se, co se týče všech vymezených předpokladů a vlastností v kapitole 2, nijak výrazně neliší od výrobků, u kterých jejich prodeje odpovídají růstové funkci, ovšem vyšší cena a tím i odlišná propagace, především zaměřená na každoroční vánoční období, neumožňuje aplikaci růstových modelů.

5.2 Ekonomický a sociální význam parametrů

Z matematického hlediska je význam parametrů jasně definován, jak je popsáno v kapitolách 1.1.2 a 1.2.2, ovšem pokud se modely aplikují na ekonomický jev, je možné

hledat jistý ekonomický či sociální význam těchto parametrů, které mohou skrývat chování trhu popřípadě chování zákazníků.

5.2.1 Parametry u Gompertzovy regrese

Koeficient β , který nabývá hodnot z intervalu (0; 1), má vliv na strmost křivky, jak již bylo uvedeno výše. Z ekonomického hlediska můžeme říci, že čím více se β blíží jedné tím je křivka méně strmá, tzn., že počáteční fáze životního cyklu trvají delší období. Z hlediska sociálního a ekonomického, tedy z pohledu zákazníka, lze říci, že za stejného předpokladu pohybu β , zákazník pomaleji reaguje na reference popřípadě na propagaci výrobku. Za těchto předpokladů můžeme koeficient β chápat jako koeficient rychlosti referencí o daném produktu.

Koeficient α , který také nabývá hodnot z intervalu (0; 1), má z ekonomického a sociálního hlediska opačný význam, než koeficient β . Zde platí, že čím více se α blíží jedné, tím rychleji prodeje dosahují maxima, tzn., že křivka je více strmá. Pokud tedy chceme chápat koeficient β jako koeficient rychlosti referencí, tzn., jak rychle zákazník reaguje na přítomnost nového výrobku na trhu, musí platit předpoklad, že koeficient α je konstantní. Tato vazba platí i opačně.

Koeficient κ určuje saturační mez, tedy nasycení trhu. Ekonomický význam spočívá v maximální velikosti poptávky, kterou je prodejce schopen uspokojit v určitém čase. Saturační mez svojí velikostí odpovídá 3. fázi životního cyklu – zralosti. Stav, kdy prodeje dosahují maxima, a neexistuje další prostor pro rozvoj výrobku.

5.2.2 Parametry u logistické regrese

Koeficient β , který, stejně jako u Gompertzovy regrese, nabývá hodnot z intervalu (0; 1), má vliv na strmost křivky. Z ekonomického hlediska můžeme, obdobně jako u Gompertzovy regrese, říci, že čím více se β blíží jedné, tím je křivka méně strmá, tzn., že počáteční fáze životního cyklu trvají delší období. Proto i sociální a ekonomický význam parametru β , můžeme definovat jako koeficient rychlosti referencí o daném produktu.

Význam koeficientu α , který nabývá hodnot větších než jedna, je v tomto případě odlišný. Z ekonomického hlediska určuje dobu, za kterou výrobek opouští zaváděcí fázi. Důležité je, že při konstantním koeficientu β se rychlost dosažení další fáze neboli strmost křivky, nemění. Můžeme tedy říci, že koeficient α určuje dobu, po jakou výrobek setrvá v počáteční fázi, nebo jinak řečeno, za jakou dobu dosáhne fáze další. Z pohledu zákazníka se jedná o dobu, kterou zákazník potřebuje k objevení nového výrobku na trhu, nejedná se tedy o ekvivalent

koeficientu rychlosti jako v případě parametru α u Gompertzovy regrese. Koeficient κ určuje stejně jako u Gompertzovy regrese saturační mez.

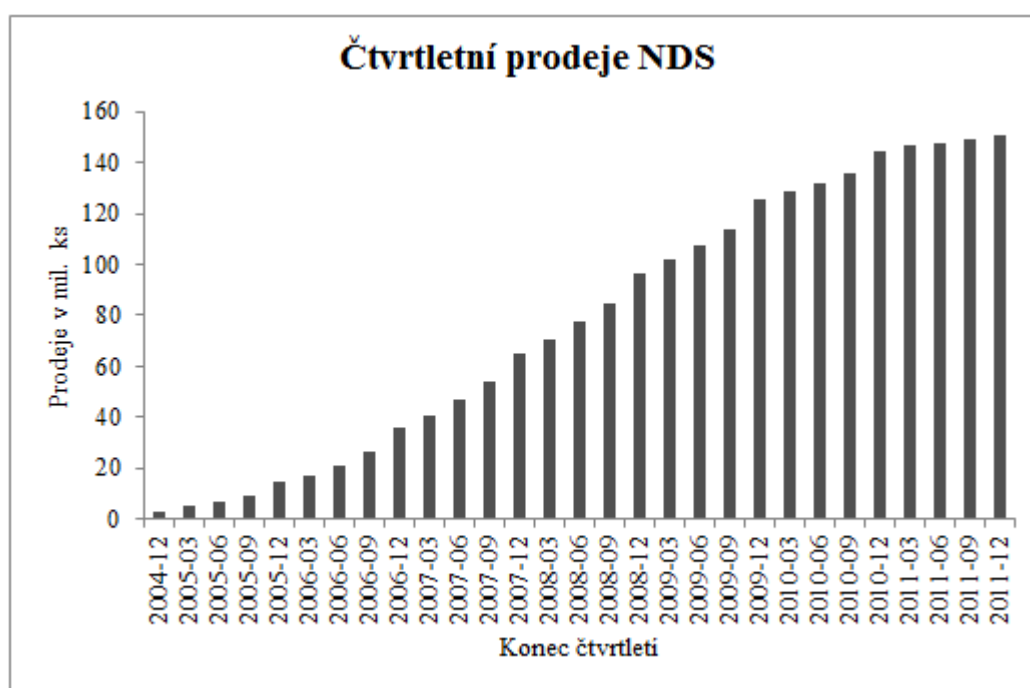
V této kapitole byly uvedeny pouze možné významy jednotlivých parametrů, pokud by se ovšem prodejci podařilo přesně analyzovat a definovat význam parametrů určitého modelu a přiřadit jim skutečné jevy, které ovlivňují trh a prodeje jeho výrobku, byl by schopen zkonstruovat predikční funkci, která by byla důležitým nástrojem v rozhodovacím a řídicím procesu firmy.

6 PRAKTICKÁ ČÁST

Obsahem této kapitoly bude aplikace růstových modelů, kdy pro odhad parametrů budou použity metody částečných součtů, dílčích průměrů a vybraných bodů, které jsou podrobněji popsány v kapitole 4. Dále bude pomocí interpolačních kritérií určena přesnost vypočtených modelů. Veškeré výpočty budou prováděny v programu MS Excel 2010. Posledním krokem bude výpočet modelu v počítačovém programu Statistica 10.

6.1 Ilustrativní příklad č. 1

V roce 2004 byla na trh uvedena herní konzole Nintendo DS (NDS), která se ihned po uvedení na trh stala velice populární. Prodeje této konzole každoročně rostou a trend, který tento vývoj vystihuje, je zobrazen na grafu č. 5.



Graf 5: Čtvrtletní prodeje NDS

Zdroj: vlastní zpracování v MS Excel, [15] až [43]

Z grafu č. 5 je zřejmé, že vývoj čtvrtletních prodejů odpovídá růstové křivce. Nejprve bude aplikován model Gompertzovy regrese za použití různých metod odhadu parametrů a poté výpočet v programu Statistica 10.

6.1.1 Odhad parametrů Gompertzovy regrese

Jelikož budeme využívat metod odhadů parametrů, jakých se používá u modifikované posunuté exponenciály, musíme si tvar Gompertzovy funkce upravit na požadovaný zápis viz

kapitola 4 $Y = \kappa \alpha^{\beta x}$. Transformací zlogaritmováním celé rovnice získáme $\ln Y = \ln \kappa + \beta^x \ln \alpha$.

Prvním krokem bude zlogaritmování získaných hodnot Y_i . Námi požadované hodnoty jsou v tabulce č. 1 v řádcích $\ln Y$.

Tabulka 1: Ilustrativní příklad č. 1 - tabulka hodnot

x_i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
y_i	3	5	7	9	14	17	21	27	36	40	47	54	65	71	78
$\ln y_i$	1,04	1,66	1,89	2,18	2,67	2,82	3,06	3,29	3,57	3,70	3,86	3,98	4,17	4,26	4,35
x_i	16	17	18	19	20	21	22	23	24	25	26	27	28	29	
y_i	84	96	102	108	113	125	129	132	136	145	146	148	149	151	
$\ln y_i$	4,43	4,57	4,62	4,68	4,73	4,83	4,86	4,88	4,91	4,97	4,99	5,00	5,00	5,02	

Zdroj: [15] až [43]

Metoda částečných součtů

Druhým krokem je výpočet parametrů α , β a κ nejprve metodou částečných součtů. Po dosazení do vzorce pro výpočet parametru b získáme hodnotu $b = \sqrt[9]{\frac{44,45 - 39,79}{39,79 - 27,03}} = 0,894$.

Pro výpočet koeficientu a dosadíme do vzorce a získáme $\ln a = \frac{(0,894 - 1)(39,79 - 27,03)}{0,894(0,894^9 - 1)^2} = -3,75$. Úpravou dostaneme hodnotu koeficientu

$a = 0,0235$. Posledním výpočtem je dosazení do vzorce pro koeficient k , kdy dostaneme

$$\ln k = \frac{1}{m} (27,03 - 3,75 \cdot 0,894 \frac{(0,894^9 - 1)}{(0,894 - 1)}) = 5,23. \text{ Úpravou dostaneme } k = 188,28. \text{ Predikční}$$

funkce získaná metodou částečných součtů má tvar: $y = 188,28 \cdot 0,0235^{0,894^x}$.

Metoda dílčích průměrů

Tato metoda se liší ve výpočtu koeficientů a a k . Pro výpočet těchto dvou koeficientů se využívá metoda nejmenších čtverců. Nejprve je však nutné vypočítat koeficient b , kdy po

dosazení do vzorce dostaneme $b = \frac{27-5}{2} \sqrt{\frac{4,49 - 4,42}{4,42 - 3}} = 0,912$. Další výpočet se provádí metodou

nejmenších čtverců kdy podle modifikované rovnice $\ln Y = \ln k + \beta^x \cdot \ln \alpha$ za Y dosadíme hodnoty $\ln Y$ a za X dosadíme hodnoty β^x . Výpočet regrese provedeme pomocí MS Excel následovně:

Karta Data – doplněk Analýza dat – nástroj Regrese

Po dosazení požadovaných hodnot získáme odhady koeficientů α a β modifikované funkce, proto je musíme upravit podle vztahů $\alpha = \ln k$ a $\beta = \ln a$. Úpravou jsme získali $a = 0,0108$ a $k = 233,74$. Predikční funkce získaná metodou dílčích průměrů má tvar:

$$y = 233,74 \cdot 0,0108^{0,912^x}$$

Metoda vybraných bodů

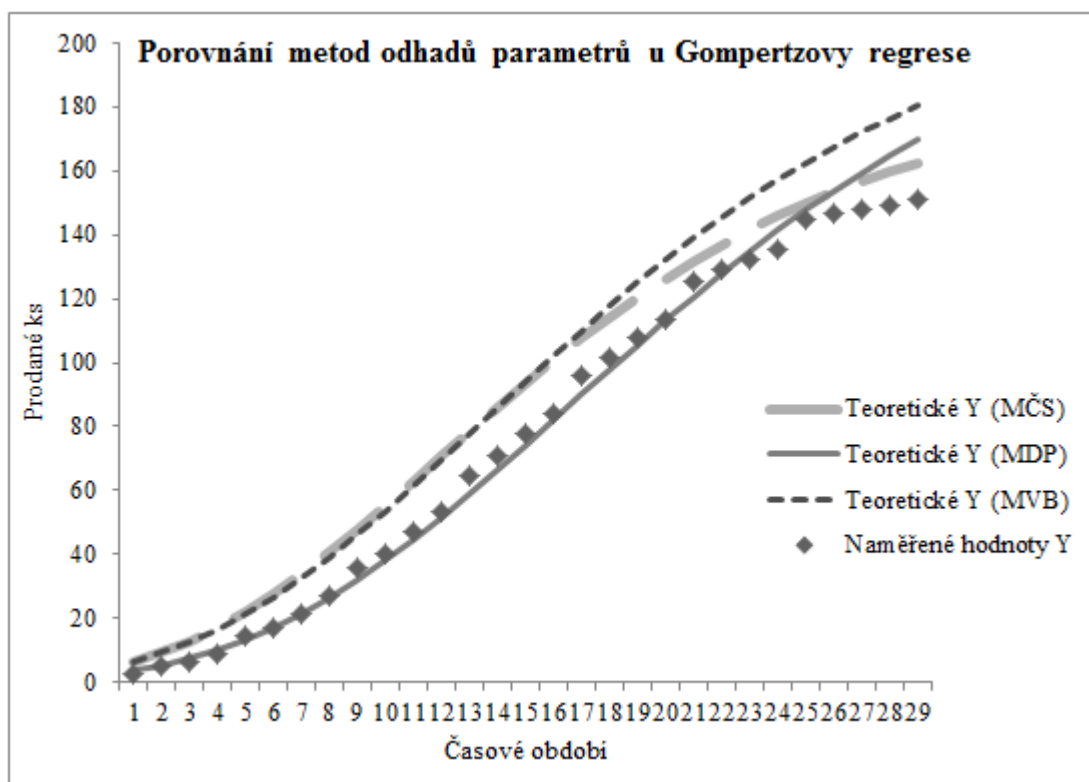
Poslední metodou je metoda vybraných bodů. Principem je taková volba hodnot x a m , kdy se x zpravidla rovná jedné a m je přirozené číslo. V našem případě $x = 1$, $m = 9$. Druhým krokem je výběr hodnot z náhodného výběru a to y_x , y_{x+m} a y_{x+2m} . Po dosazení do rovnic získáme hodnoty koeficientů:

$$b = \sqrt[9]{\frac{4,82 - 3,98}{3,98 - 1,89}} = 0,905,$$

$$\ln a = \frac{3,98 - 1,89}{0,905^1 (0,905^9 - 1)} = -3,88, \text{ po úpravě } a = 0,0206$$

$$\ln k = 1,89 - (-3,88) \cdot 0,905^1 = 5,407, \text{ po úpravě } k = 223,12$$

Metodou vybraných bodů jsme získali predikční rovnici v tomto tvaru:
 $y = 223,12 \cdot 0,0206^{0,905^x}$. Na následujícím grafu je možné porovnat jednotlivé metody a jejich vypočítané predikční funkce.



Graf 6: Porovnání metod odhadů parametrů u GR

Zdroj: vlastní zpracování MS Excel

Z grafu č. 6 je patrné, že nejlepších výsledků bylo dosaženo pomocí metody dílčích průměrů. Pro ověření budou metody a jejich výsledné funkce porovnány pomocí interpolačních kritérií a to konkrétně střední čtvercovou chybou a střední absolutní chybou odhadu. Porovnání je v tabulce níže, kde je možné vidět, že nejvhodnější predikční funkce bylo docíleno metodou dílčích průměrů (MDP). Jednotlivá interpolační kritéria dosáhly následujících výsledků:

$$\text{MSE} = 35,7$$

$$\text{MAE} = 3,94$$

Pomocí metody částečných součtů bylo dosaženo těchto výsledků interpolačních kritérií:

$$\text{MSE} = 132,44$$

$$\text{MAE} = 10,89$$

A nejhorsích výsledků bylo dosaženo metodou vybraných bodů a to konkrétně:

$$\text{MSE} = 264,06$$

$$\text{MAE} = 15,02$$

Tabulka 2: Interpoláčn kritria . 1 - ilustrativn pklad . 1

X	Y	Metody odhad paramet			Mean Square Error			Mean Absolute Error		
		Y(MS)	Y(MDP)	Y(MVB)	(MS)	(MDP)	(MVB)	(MS)	(MDP)	(MVB)
1	2,84	6,58	3,76	6,65	13,95	0,85	14,52	3,74	0,92	3,81
2	5,27	9,38	5,40	9,30	16,89	0,02	16,21	4,11	0,13	4,03
3	6,65	12,89	7,51	12,59	38,89	0,73	35,26	6,24	0,86	5,94
4	8,83	17,12	10,14	16,56	68,68	1,72	59,73	8,29	1,31	7,73
5	14,43	22,07	13,35	21,22	58,30	1,17	46,10	7,64	1,08	6,79
6	16,73	27,69	17,15	26,56	120,09	0,17	96,58	10,96	0,42	9,83
7	21,27	33,92	21,55	32,53	160,02	0,08	126,89	12,65	0,28	11,26
8	26,82	40,67	26,55	39,09	191,81	0,07	150,60	13,85	0,27	12,27
9	35,61	47,83	32,12	46,16	149,44	12,21	111,22	12,22	3,49	10,55
10	40,29	55,30	38,21	53,64	225,40	4,34	178,22	15,01	2,08	13,35
11	47,27	62,96	44,77	61,45	246,29	6,25	201,09	15,69	2,50	14,18
12	53,64	70,71	51,74	69,49	291,29	3,63	251,27	17,07	1,90	15,85
13	64,79	78,43	59,03	77,67	186,15	33,13	165,86	13,64	5,76	12,88
14	70,60	86,05	66,59	85,89	238,83	16,09	233,84	15,45	4,01	15,29
15	77,54	93,49	74,32	94,08	254,49	10,35	273,51	15,95	3,22	16,54
16	84,33	100,69	82,16	102,15	267,53	4,70	317,70	16,36	2,17	17,82
17	96,22	107,59	90,03	110,06	129,18	38,28	191,42	11,37	6,19	13,84
18	101,78	114,15	97,87	117,73	153,12	15,27	254,36	12,37	3,91	15,95
19	107,75	120,37	105,62	125,13	159,16	4,54	302,06	12,62	2,13	17,38
20	113,48	126,21	113,22	132,23	161,93	0,07	351,39	12,73	0,26	18,75
21	125,13	131,67	120,64	138,99	42,72	20,19	192,11	6,54	4,49	13,86
22	128,89	136,75	127,83	145,41	61,75	1,13	272,83	7,86	1,06	16,52
23	132,04	141,46	134,76	151,47	88,70	7,39	377,44	9,42	2,72	19,43
24	135,58	145,81	141,41	157,17	104,59	34,01	466,02	10,23	5,83	21,59
25	144,59	149,81	147,77	162,51	27,23	10,11	321,05	5,22	3,18	17,92
26	146,42	153,48	153,82	167,50	49,82	54,76	444,17	7,06	7,40	21,08
27	147,86	156,84	159,56	172,14	80,57	136,82	589,47	8,98	11,70	24,28
28	149,00	159,90	164,98	176,45	118,82	255,30	753,53	10,90	15,98	27,45
29	151,06	162,69	170,09	180,44	135,28	361,98	863,40	11,63	19,03	29,38
Suma					3840,88	1035,37	7657,87	315,77	114,28	435,53
MSE / MAE					132,44	35,70	264,06	10,89	3,94	15,02
MIN MSE / MAE						35,70			3,94	

Optimalizace výsledku Gompertzovy regrese pomocí řešitele

Výsledné hodnoty parametrů, vypočtené jakoukoli metodou, můžeme dále optimalizovat pomocí modulu řešitel v MS Excel. Výsledky, dosažené tímto postupem, jsou často téměř identické s výsledky dosaženými výpočtem v programu Statistica 10.

Řešitel je součástí sady příkazů a jeho podstatou je proces, při kterém se mění hodnoty v buňkách za účelem posouzení vlivu změn na výsledky vzorců v listu. Pomocí řešitele je možné najít optimální (maximální nebo minimální) hodnotu pro vzorec s ohledem na omezující podmínky (neboli limity) určené hodnotami jiných buněk vzorce v listu. Řešitel pracuje se skupinou buněk označovaných jako rozhodovací proměnné nebo proměnné buňky, které se podílejí na výpočtu vzorců v buňce cíle a buňkách s omezujícími podmínkami. Řešitel upravuje hodnoty v buňkách rozhodovacích proměnných tak, aby byly splněny limity buněk s omezujícími podmínkami a byl nalezen výsledek požadovaný pro cílovou buňku. Řešitel projde takovým počtem iterací, dokud není dosaženo optimální řešení. [1]

Na následující tabulce bude popsán postup v MS Excel.

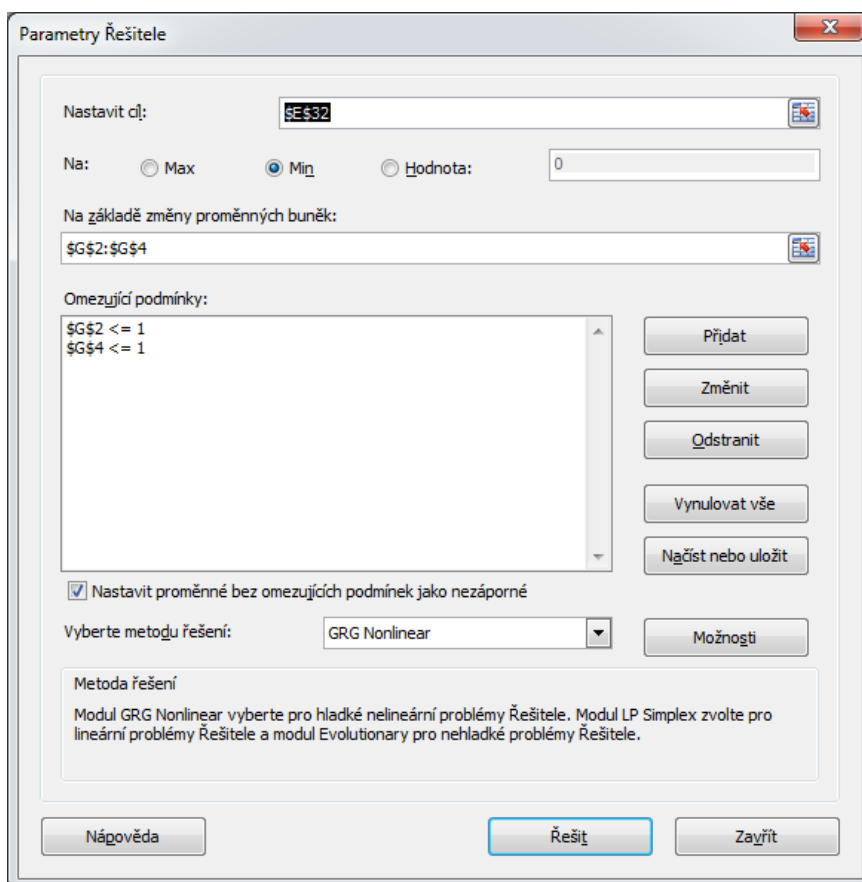
Tabulka 3: Optimalizace Gompertzovy regrese

X	Y	ln Y	Y teor.	Čtverce	b	0,88749
1	2,84	1,043804	2,314728	0,275911	k	180,337
2	5,27	1,66203	3,778492	2,224596	a	0,00739
3	6,65	1,894617	5,837049	0,66089		
4	8,83	2,178155	8,586539	0,059273		
5	14,43	2,669309	12,09439	5,455086		
6	16,73	2,817204	16,39126	0,114747		
7	21,27	3,057298	21,46774	0,039102		
8	26,82	3,289148	27,2758	0,207752		
9	35,61	3,572626	33,73407	3,519116		
10	40,29	3,696103	40,73582	0,198757		
11	47,27	3,855876	48,15806	0,788645		
12	53,64	3,982295	55,87055	4,975351		
13	64,79	4,171151	63,74389	1,094339		
14	70,6	4,25703	71,65599	1,115118		
15	77,54	4,350794	79,49677	3,828931		
16	84,33	4,434738	87,17112	8,071977		
17	96,22	4,566637	94,60038	2,623163		
18	101,78	4,622814	101,7225	0,003312		
19	107,75	4,679814	108,4911	0,549212		
20	113,48	4,731627	114,8745	1,944637		
21	125,13	4,829353	120,8536	18,28736		
22	128,89	4,858959	126,4202	6,099858		
23	132,04	4,883105	131,5749	0,216324		
24	135,58	4,909562	136,3254	0,55563		
25	144,59	4,973902	140,6849	15,24972		
26	146,42	4,986479	144,6706	3,060523		
27	147,86	4,996266	148,3023	0,195607		
28	149	5,003946	151,6017	6,768718		
29	151,06	5,017677	154,5913	12,47007		
				Součet čtverců		100,654

Řešitel spustíme příkazem *Data – Řešitel*. Samotný modul obsahuje 5 argumentů. Nejprve nastavíme buňku, která má být minimalizována nebo maximalizována. V našem případě budeme minimalizovat součet čtverců. Nastavíme tedy buňku, která obsahuje součet čtverců. Druhým krokem je nastavení, zda budeme uvažovat minimalizační funkci nebo maximalizační. Třetím krokem je nastavení buněk, které může řešitel měnit, takovým způsobem, aby dosáhl co nejlepšího výsledku. Nastavíme tedy 3 buňky, které obsahují parametry *a*, *b*, *k*. Je potřeba uvést přibližné hodnoty těchto parametrů, nejlépe výsledky dosažené jednou z předešlých metod. Dále nastavíme omezující podmínky, a to konkrétně pro

parametry a a b , kdy jsou menší než 1. Posledním krokem je výběr metody řešení, kterou je v našem případě metoda *GRC Nonlinear*.

Nastavení řešitele pro daný příklad je zobrazen na obrázku č. 2.

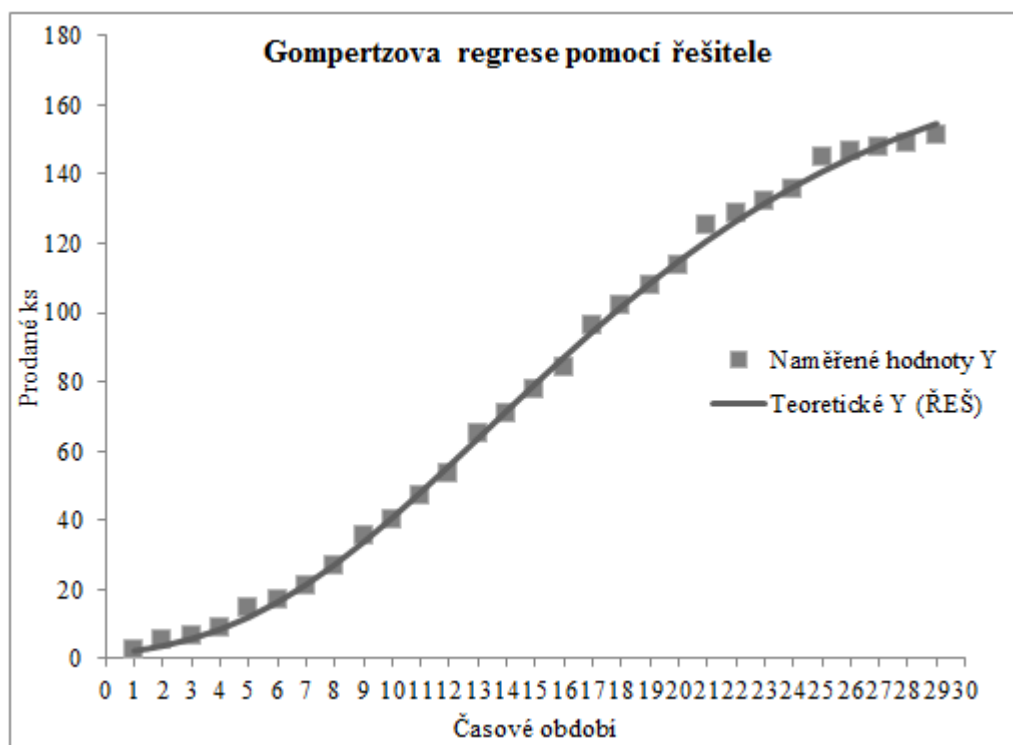


Obrázek 2: Nastavení řešitele

Zdroj: výstup z MS Excel

Hodnoty parametrů jsou již zobrazeny v tabulce č. 3. Funkce získaná po optimalizaci parametrů pomocí řešitele má následující tvar: $y = 180,337 \cdot 0,00739^{0,88749}$.

Gompertzova křivka získaná optimalizací pomocí řešitele je zobrazena na následujícím grafu.

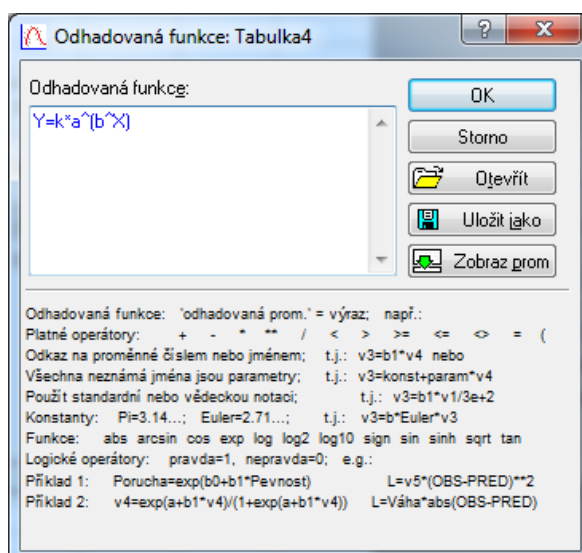


Graf 7: GR pomocí řešitele

Zdroj: vlastní zpracování v MS Excel

Výpočet Gompertzovy regrese v programu Statistica 10

Posledním způsobem aplikace Gompertzovy regrese na daná data je výpočet v programu Statistica 10. Výpočet provedeme zadáním příkazu *Statistiky – Pokročilé modely – Nelineární odhad*. V dalším kroku vybereme způsob výpočtu metodu nejmenších čtverců. Dále musíme ručně zadat syntaxi odhadované funkce. V našem případě rovnici Gompertzovy křivky. Zápis je ukázán na následujícím obrázku.



Obrázek 3: Zápís Gompertzovy rovnice

Zdroj: výstup z programu Statistica 10

V posledním kroku musíme vyplnit počáteční hodnoty parametrů a hladinu významnosti α . Také máme možnost vybrat mezi dvěma metodami odhadu a to Levenberg-Marquardt nebo Gauss-Newtonovou. Pro náš výpočet zvolíme prvně jmenovanou metodu. Po provedení výpočtu programem Statistica 10, získáme následující výstupy:

Tabulka 4: Statistica 10 - výstup č. 1

Model je: $Y=k*a^{(b^X)}$ (Tabulka4)						
Záv.prom.:Y						
Hladina spolehlivosti:95.0% (alfa =0.050)						
	Odhad	Standard chyba	t-hodn. sv = 26	p-hodn.	Dol. sp. Mez	Hor. sp. Mez
k	180,3376	2,962880	60,8656	0,000000	174,2470	186,4276
a	0,0074	0,001102	6,7064	0,000000	0,0051	0,0097
b	0,8875	0,003035	292,0460	0,000000	0,8812	0,8937

Zdroj: výstup z programu Statistica 10

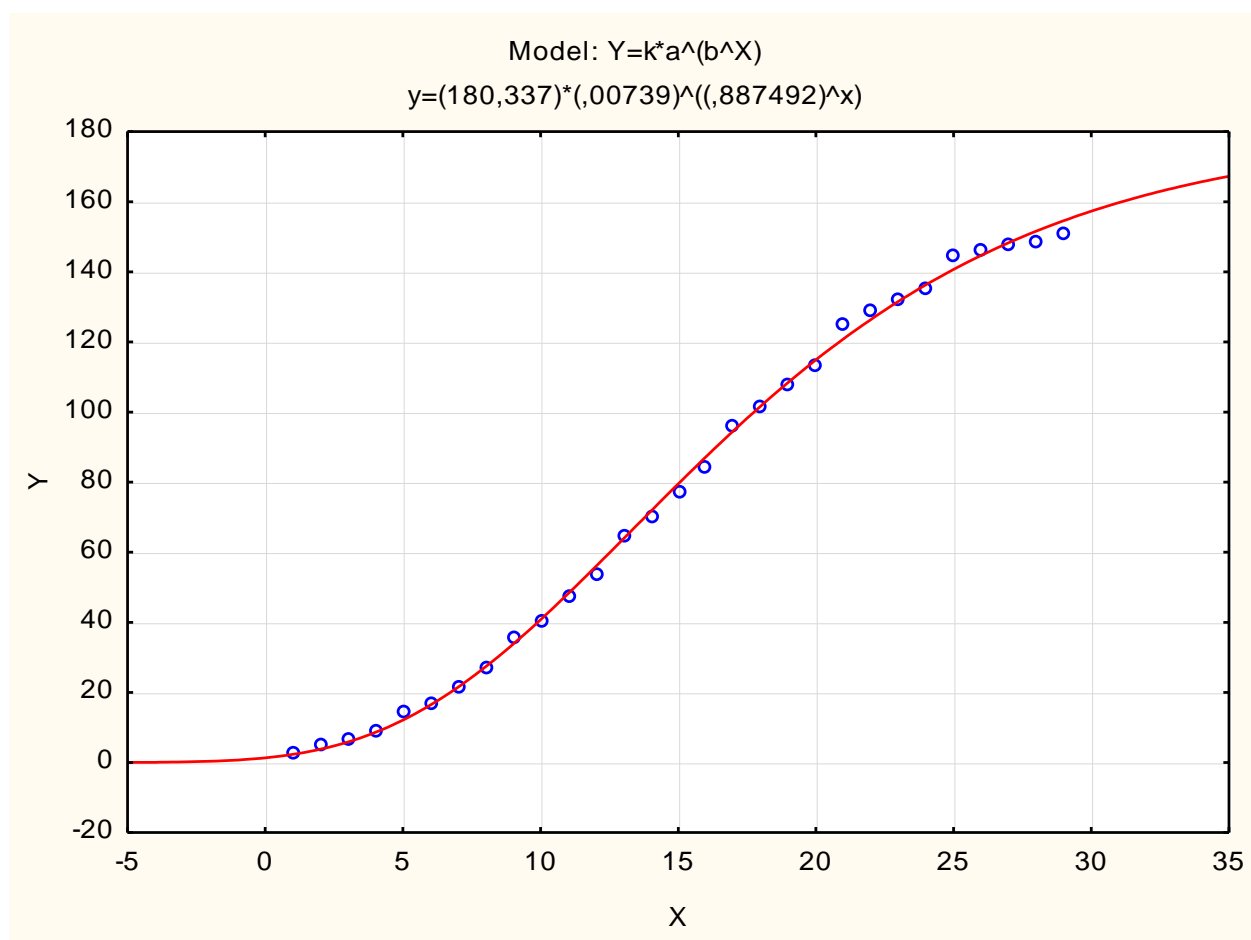
V tabulce č. 4 nalezneme v prvním sloupci Odhad konečné hodnoty jednotlivých parametrů, které jsou totožné s hodnotami získanými při optimalizaci pomocí řešitele v MS Excel. Sloupce Dol. sp. Mez a Hor. sp. Mez určují intervaly spolehlivosti pro jednotlivé parametry. Na jejich základě můžeme tvrdit, že s pravděpodobností 95% budou hodnoty parametrů modelu v následujících intervalech, kdy κ je z intervalu $\langle 174,247; 186,427 \rangle$, parametr α je z intervalu $\langle 0,0051; 0,0097 \rangle$ a parametr β je z intervalu $\langle 0,881; 0,894 \rangle$.

Tabulka 5: Statistica 10 - výstup č. 2

Model je: $Y=k*a^{(b^X)}$ (Tabulka4)				
Záv.prom.:Y				
	Ztrátová funkce	k	a	b
1	87,63748	175,0000	0,006000	0,860000
2	22,32994	174,6919	0,009179	0,890511
3	10,08052	180,3704	0,007253	0,887512
4	10,03269	180,3154	0,007379	0,887469
5	10,03263	180,3379	0,007389	0,887492
6	10,03263	180,3379	0,007390	0,887492

Zdroj: výstup z programu Statistica 10

Tabulka č. 5 zobrazuje jednotlivé iterace, kterými program Statistica 10 procházel, než našel optimální hodnoty parametrů. V prvním řádku jsou zobrazeny uživatelem nastavené počáteční hodnoty. V posledním řádku jsou konečné hodnoty parametrů. Počet iterací záleží na zadání počátečních parametrů. Čím blíže budou počáteční hodnoty optimálním hodnotám, tím menší bude počet iterací. Na grafu č. 8 je zobrazena Gompertzova křivka, která prolíná naměřené hodnoty y_i .



Graf 8: Statistica 10 - výstup č. 3

Zdroj: Zdroj: výstup z programu Statistica 10

Dále je na výstupu č. 3 uvedena rovnice v zápisu $y=(180,337)*(,00739)^{((,88749)^x)}$, který je ekvivalentem zápisu $y = 180,337 \cdot 0,00739^{0,88749^x}$.

Jak přesně křivka prolíná naměřené hodnoty, určuje index determinace R^2 . Tento ukazatel je také součástí výstupu při výpočtu programem Statistica 10. Index determinace nebo také podíl rozptylu vysvětlený modelem je roven hodnotě 0,9987, hodnota $r = 0,999$.

V poslední části ilustrativního příkladu č. 1 bude pomocí interpolačních kritérií porovnán výpočet v programu Statistica 10 a výpočet dosažený optimalizací řešitelem. V následující tabulce budeme používat zkratky ŘEŠ pro řešitele a STA pro program Statistica 10.

Tabulka 6: Interpolační kritéria č. 2 - ilustrativní příklad č. 1

X	Y	Metody odhadu		MSE		MAE	
		Y(ŘEŠ)	Y(STA)	(ŘEŠ)	(STA)	(ŘEŠ)	(STA)
1	2,84	2,31	2,31	0,28	0,28	0,53	0,53
2	5,27	3,78	3,78	2,22	2,22	1,49	1,49
3	6,65	5,84	5,84	0,66	0,66	0,81	0,81
4	8,83	8,59	8,59	0,06	0,06	0,24	0,24
5	14,43	12,09	12,09	5,46	5,46	2,34	2,34
6	16,73	16,39	16,39	0,11	0,11	0,34	0,34
7	21,27	21,47	21,47	0,04	0,04	0,20	0,20
8	26,82	27,28	27,28	0,21	0,21	0,46	0,46
9	35,61	33,73	33,73	3,52	3,52	1,88	1,88
10	40,29	40,74	40,74	0,20	0,20	0,45	0,45
11	47,27	48,16	48,16	0,79	0,79	0,89	0,89
12	53,64	55,87	55,87	4,98	4,98	2,23	2,23
13	64,79	63,74	63,74	1,09	1,09	1,05	1,05
14	70,60	71,66	71,66	1,12	1,12	1,06	1,06
15	77,54	79,50	79,50	3,83	3,83	1,96	1,96
16	84,33	87,17	87,17	8,07	8,07	2,84	2,84
17	96,22	94,60	94,60	2,62	2,62	1,62	1,62
18	101,78	101,72	101,72	0,00	0,00	0,06	0,06
19	107,75	108,49	108,49	0,55	0,55	0,74	0,74
20	113,48	114,87	114,87	1,94	1,94	1,39	1,39
21	125,13	120,85	120,85	18,29	18,29	4,28	4,28
22	128,89	126,42	126,42	6,10	6,10	2,47	2,47
23	132,04	131,57	131,57	0,22	0,22	0,47	0,47
24	135,58	136,33	136,33	0,56	0,56	0,75	0,75
25	144,59	140,68	140,68	15,25	15,25	3,91	3,91
26	146,42	144,67	144,67	3,06	3,06	1,75	1,75
27	147,86	148,30	148,30	0,20	0,20	0,44	0,44
28	149,00	151,60	151,60	6,77	6,77	2,60	2,60
29	151,06	154,59	154,59	12,47	12,47	3,53	3,53
Suma				100,65	100,65	42,74	42,74
MSE / MAE				3,47	3,47	1,47	1,47
MIN MSE / MAE				3,47	3,47	1,47	1,47

Zdroj: vlastní zpracování v MS Excel

Jak můžeme sledovat v porovnání obou metod, výsledky dosažené optimalizací pomocí řešitele v MS Excel jsou identické s výsledky dosaženými výpočtem v programu Statistica 10. V porovnání s metodou dílčích průměrů, která dosáhla v předchozí části nejlepších výsledků, dosahuje výpočet řešitelem a programem Statistica 10 ještě přesnějších výsledků.

$$\text{MSE (ŘEŠ, STA)} = 3,47; \text{MSE (MDP)} = 35,7$$

$$\text{MAE}(\text{ŘEŠ}, \text{STA}) = 1,47; \text{MAE}(\text{MDP}) = 3,94$$

Závěrem lze tedy říci, že výpočet pomocí statistického programu Statistica 10 dává nejpřesnější výsledky při modelování nelineární Gompertzovy regrese. Kvalitních výsledků můžeme dosáhnout i pro běžného uživatele dostupnějším programem MS Excel. Ovšem výpočet je komplikovanější jelikož se skládá ze samotného odhadu parametrů a z následné optimalizace pomocí modulu řešitel. V našem příkladu bylo dosaženo pomocí obou metod stejných výsledků.

6.1.2 Predikce budoucích hodnot pomocí Gompertzovy rovnice

Pro odhad budoucích prodejů použijeme rovnici z výpočtu v programu Statistica 10, jelikož, zde bylo dosaženo nejlepších odhadů parametrů. Naše predikční rovnice má tvar: $Y = 180,337 \cdot 0,00739^{0,88749^x}$. Připomene, že v našem příkladu naměřené hodnoty Y představují čtvrtletní prodeje herní konzole NDS. K dispozici jsou skutečné hodnoty za 29 období (4. čtvrtletí 2004 až 4. čtvrtletí 2011), proto provedeme odhad hodnot pro dvě následující období (1. a 2. čtvrtletí 2012). Samotný výpočet je jednoduchý, kdy do rovnice dosadíme požadované x_i , v našem případě hodnoty 30 a 31. Predikce budoucích hodnot je následující:

$$y_{30} = 180,337 \cdot 0,00739^{0,88749^{30}} = 157,29$$

$$y_{31} = 180,337 \cdot 0,00739^{0,88749^{31}} = 159,73$$

Můžeme tedy říci, že na základě predikční rovnice, získané výpočtem v programu Statistica 10 je odhad prodejů herní konzole NDS pro 1. čtvrtletí roku 2012 157 290 000 kusů a pro 2. čtvrtletí roku 2012 je odhad prodejů 159 730 000 kusů.

6.1.3 Extrapolace dat u Gompertzovy regrese

Podrobný popis této metody je uveden v kapitole 3.3.2. Pro náš příklad vynecháme poslední dvě pozorování a získáme následující náhodný výběr:

Tabulka 7: Extrapolace dat - náhodný výběr č. 1

x_i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
y_i	3	5	7	9	14	17	21	27	36	40	47	54	65	71	78
$\ln y_i$	1,04	1,66	1,89	2,18	2,67	2,82	3,06	3,29	3,57	3,70	3,86	3,98	4,17	4,26	4,35
x_i	16	17	18	19	20	21	22	23	24	25	26	27			
y_i	84	96	102	108	113	125	129	132	136	145	146	148			
$\ln y_i$	4,43	4,57	4,62	4,68	4,73	4,83	4,86	4,88	4,91	4,97	4,99	5,00			

Zdroj: [15] až [43]

Výpočtem parametrů predikční rovnice provedeným v programu Statistica 10 bylo dosaženo těchto hodnot: $a = 0,0085$, $b = 0,8925$ a $k = 187,7302$. Predikční rovnice má tedy tvar $Y = 187,7302 \cdot 0,0085^{0,8925^x}$.

Odhad budoucích hodnot je následující:

$$y_{28} = 187,7302 \cdot 0,0085^{0,8925^{28}} = 154,10$$

$$y_{29} = 187,7302 \cdot 0,0085^{0,8925^{29}} = 157,41$$

Pro zhodnocení simulace použijeme Theilův koeficient, kdy po dosazení získáme:

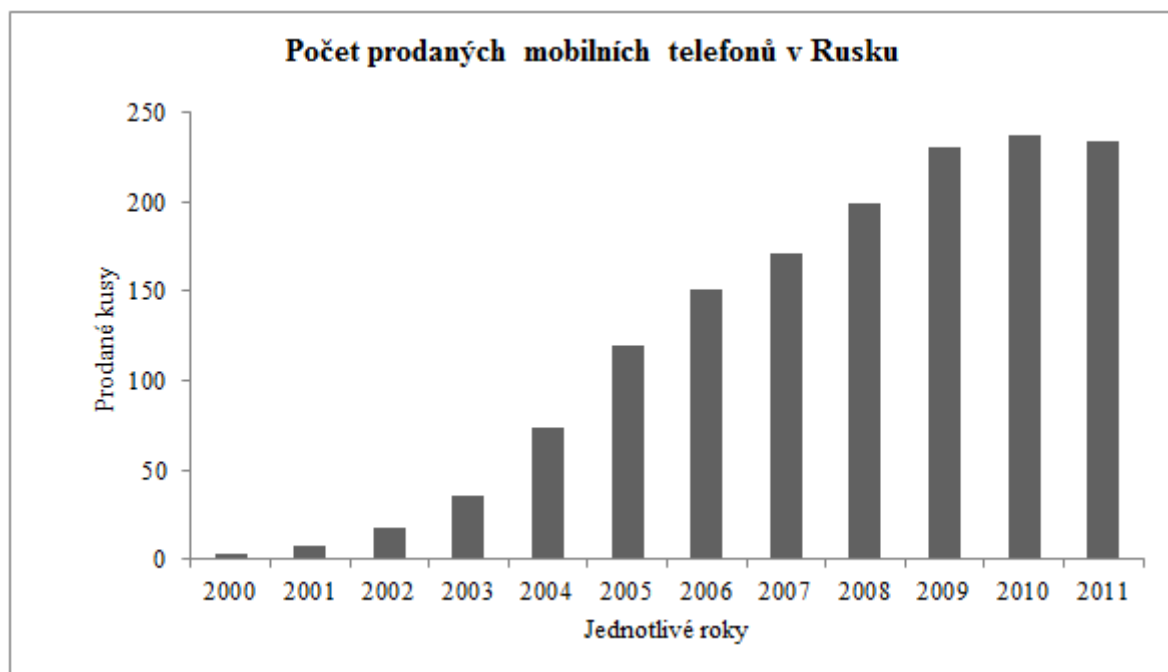
$$T_H^2 = \frac{(149 - 154,10)^2 + (151,06 - 157,41)^2}{149^2 + 151,06^2} = 0,00147$$

Na základě výsledné hodnoty Theilova koeficientu můžeme říci, že simulace dosahuje velmi dobrých výsledků. V absolutních hodnotách došlo u 28. období k odchylce 5,1 a u 29. období k odchylce 6,41.

6.2 Ilustrativní příklad č. 2

Na následující časovou řadu bude aplikována logistická regrese, kdy parametry křivky budou odhadnuty stejnými způsoby jako v ilustrativním příkladu č. 1.

Rozvoj informačních technologií je nedílnou součástí současného světového vývoje. Důkazem je i obrovský rozvoj telekomunikačních technologií ke konci 20. století. Následující graf zobrazuje počet prodaných mobilních telefonů v Rusku od roku 2000.



Graf 9: Počet prodaných mobilních telefonů v Rusku

Zdroj: vlastní zpracování v MS Excel, [13]

Stejně jako v předchozím příkladu, i zde je patrný růstový trend vývoje počtu prodaných kusů daného produktu.

6.2.1 Odhad parametrů logistické regrese

Provedeme nejprve úpravu na požadovaný zápis viz kapitola 4. Transformaci provedeme

převrácením výrazu $Y = \frac{k}{1 + \alpha\beta^x}$ a obdržíme:

$$\frac{1}{Y} = \frac{1}{k} + \frac{\alpha}{k} \beta^x$$

Naměřené hodnoty Y si upravíme na požadovaný tvar, který je ve třetím řádku tabulky níže.

Tabulka 8: Ilustrativní příklad č. 2 - tabulka hodnot

x_i	1	2	3	4	5	6	7	8	9	10	11	12
y_i	3,26	7,75	17,61	36,14	73,72	120,00	150,67	171,20	199,52	230,50	237,69	234,40
$1/y_i$	0,30645	0,12902	0,05679	0,02767	0,01356	0,00833	0,00664	0,00584	0,00501	0,00434	0,00421	0,00427

Zdroj: [13]

Metoda částečných součtů

Po dosazení do vzorce pro odhad parametru β získáme hodnotu

$$b = \sqrt[4]{\frac{0,0178 - 0,0344}{0,0344 - 0,5199}} = 0,429. \text{ Pro odhad koeficientu } \alpha \text{ dosadíme do vzorce a získáme}$$

$$\frac{a}{k} = \frac{(0,429 - 1)(0,0344 - 0,5199)}{0,429(0,429^4 - 1)^2} = 0,691. \text{ Vynásobíme koeficientem } k \text{ a získáme hodnotu}$$

koeficientu $a = 160,27$. Posledním výpočtem je dosazení do vzorce pro koeficient k , kdy

$$\text{dostaneme } \frac{1}{k} = \frac{1}{4} (0,5199 - 0,691 \cdot 0,429 \frac{(0,429^4 - 1)}{(0,429 - 1)}) = 0,0043. \text{ Hodnotu převrátíme}$$

a dostaneme $k = 232,024$. Predikční funkce získaná metodou částečných součtů má tvar:

$$y = \frac{232,024}{1 + 160,27 \cdot 0,429^x}.$$

Metoda dílčích průměrů

Nejprve dosadíme do vzorce pro výpočet $b = \sqrt[2]{\frac{0,0046 - 0,0086}{0,0086 - 0,13}} = 0,381$. Další výpočet

se provádí metodou nejmenších čtverců, kdy za Y dosadíme hodnoty $1/Y$ a za X dosadíme hodnoty β^X . Výpočet regrese provedeme pomocí MS Excel následovně:

Karta Data – doplněk Analýza dat – nástroj Regrese

Po dosazení požadovaných hodnot získáme odhady koeficientů α a β modifikované funkce, proto je musíme upravit podle vztahů $\alpha = 1/k$ a $\beta = \alpha/k$. Poté získáme $a = 115,88$ a $k = 145,558$. Predikční funkce získaná metodou dílčích průměrů má tvar:

$$y = \frac{145,558}{1 + 115,88 \cdot 0,380^x}.$$

Metoda vybraných bodů

Po dosazení do rovnic získáme hodnoty koeficientů:

$$b = \sqrt[5]{\frac{0,0042 - 0,0083}{0,0083 - 0,306}} = 0,425$$

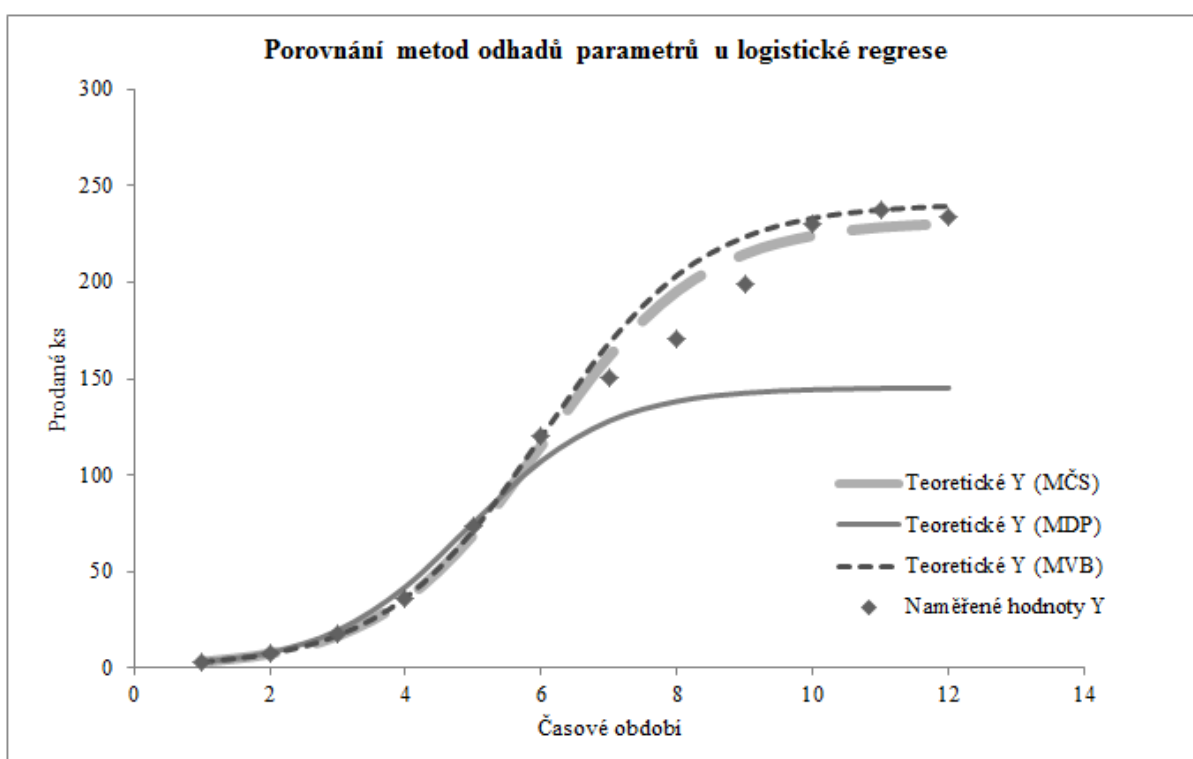
$$\frac{a}{k} = \frac{0,0083 - 0,306}{0,425(0,425^5 - 1)} = 0,712, \text{ po úpravě } a = 171,5$$

$$\frac{1}{k} = 0,306 - 0,712 \cdot 0,425 = 0,00415 \text{ a po úpravě } k = 241,007$$

Metodou vybraných bodů jsme získali predikční rovnici v tomto tvaru:

$$y = \frac{241,007}{1 + 171,5 \cdot 0,425^x}$$

Na následujícím grafu je možné porovnat jednotlivé metody a jejich vypočítané predikční funkce.



Graf 10: Porovnání metod odhadů parametrů u logistické regrese

Zdroj: vlastní zpracování v MS Excel

Na grafu č. 10 můžeme sledovat, jak odhadem parametrů metodou dílčích průměrů (MDP) bylo dosaženo nejméně kvalitní predikční funkce. Ovšem pomocí zbylých dvou metod bylo dosaženo dobrých výsledků. Pro přesnější porovnání slouží tabulka č. 9, která obsahuje interpolační kritéria pro ilustrativní příklad č. 2.

Tabulka 9: Interpolační kritéria č. 3 - ilustrativní příklad č. 2

X	Y	Metody odhadů parametrů			Mean Square Error			Mean Absolute Error		
		Y(MČS)	Y(MDP)	Y(MVB)	Y(MČS)	Y(MDP)	Y(MVB)	Y(MČS)	Y(MDP)	Y(MVB)
1	3,26	3,32	3,23	3,26	0,00	0,00	0,00	0,06	0,04	0,00
2	7,75	7,58	8,17	7,54	0,03	0,18	0,04	0,17	0,42	0,21
3	17,61	16,92	19,67	17,03	0,48	4,24	0,33	0,69	2,06	0,58
4	36,14	35,90	42,34	36,59	0,06	38,50	0,21	0,24	6,21	0,45
5	73,72	69,31	75,48	71,44	19,46	3,09	5,21	4,41	1,76	2,28
6	120,00	115,51	107,54	120,00	20,18	155,36	0,00	4,49	12,46	0,00
7	150,67	161,87	128,28	168,72	125,25	501,31	325,79	11,19	22,39	18,05
8	171,20	195,60	138,46	203,90	595,15	1072,03	1069,01	24,40	32,74	32,70
9	199,52	214,83	142,77	223,71	234,38	3220,79	584,95	15,31	56,75	24,19
10	230,50	224,31	144,48	233,34	38,31	7398,73	8,07	6,19	86,02	2,84
11	237,69	228,65	145,15	237,69	81,79	8563,94	0,00	9,04	92,54	0,00
12	234,40	230,56	145,40	239,59	14,74	7920,65	26,89	3,84	89,00	5,19
Suma					1129,83	28878,80	2020,51	80,03	402,39	86,48
MSE / MAE					94,15	2406,57	168,38	6,67	33,53	7,21
MIN MSE / MAE					94,15			6,67		

Zdroj: vlastní zpracování v MS Excel

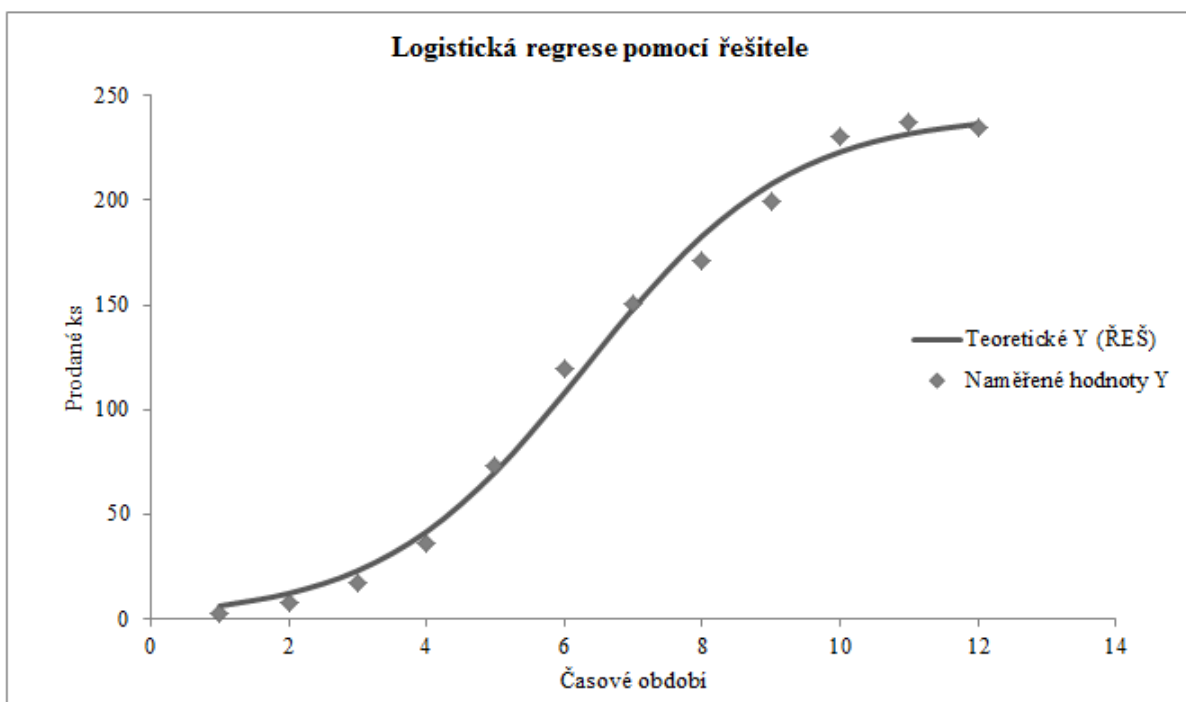
Dle výsledků interpolačních kritérií můžeme říci, že nejvhodnější predikční funkce bylo dosaženo metodou částečných součtů, kdy MSE = 94,15 a MAE = 6,67.

Optimalizace výsledku logistické regrese pomocí řešitele

Podobně jako v předchozím příkladu i zde provedeme optimalizaci hodnot parametrů v MS Excel. Stejným postupem obdržíme optimalizované hodnoty parametrů a následně

$$\text{i funkci v tomto tvaru: } y = \frac{241,5687}{1 + 71,00535 \cdot 0,508897^x}$$

Logistická křivka získaná optimalizací pomocí řešitele je zobrazena na následujícím grafu.

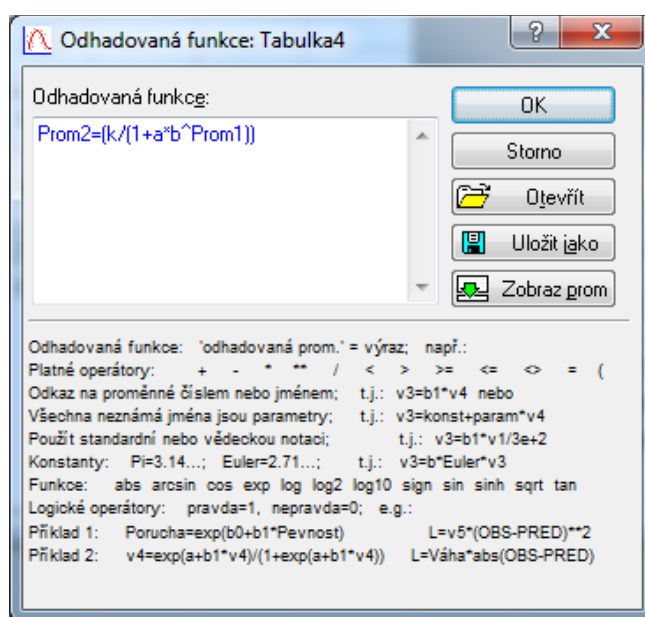


Graf 11: Logistická regrese pomocí řešitele

Zdroj: vlastní zpracování v MS Excel

Výpočet logistické regrese v programu Statistica 10

Výpočet v programu Statistica 10 je totožný jako v předchozím příkladu. S tím rozdílem, že použijeme syntaxi logistické rovnice. Zápis je ukázán na následujícím obrázku.



Obrázek 4: Zápis logistické regrese

Zdroj: výstup z programu Statistica 10

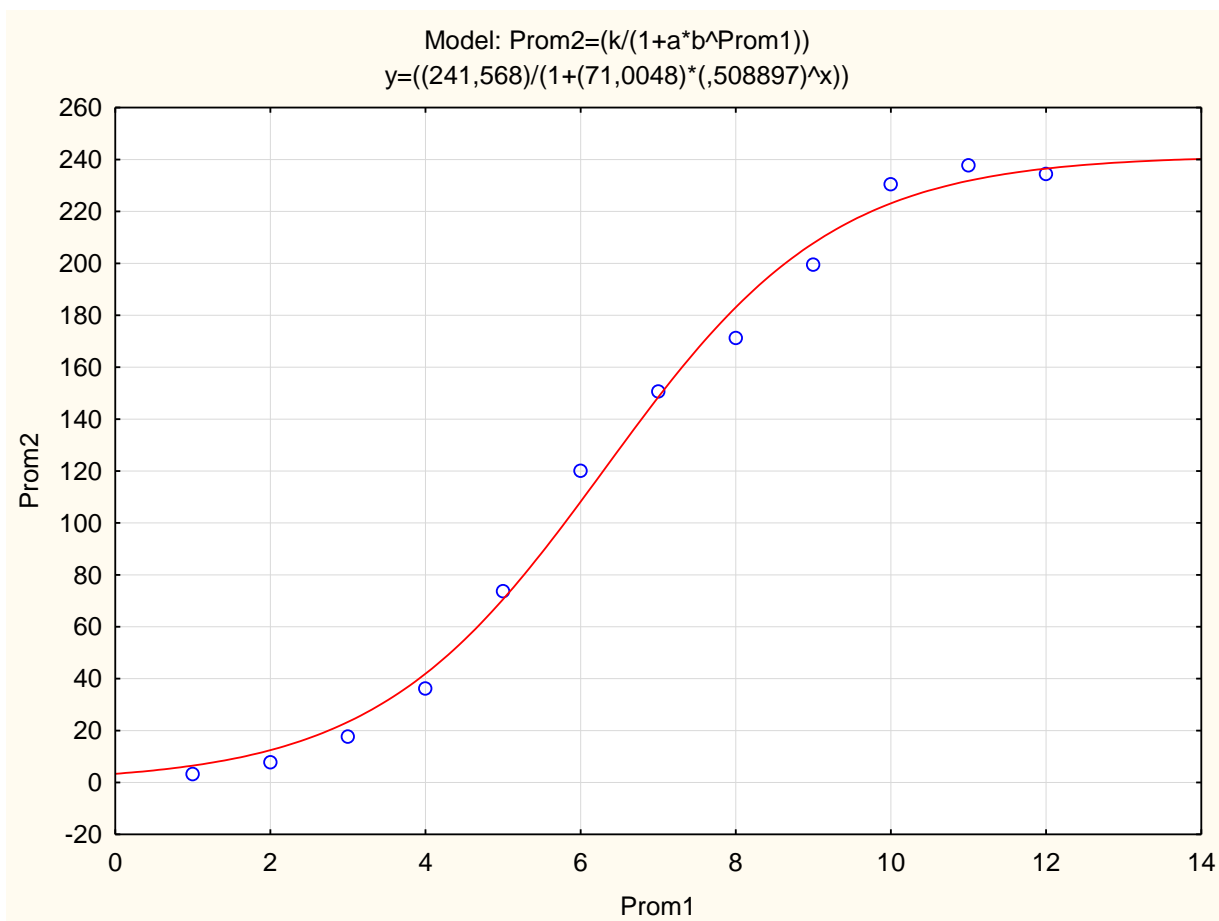
Vypočtené hodnoty parametrů jsou zobrazeny v tabulce č. 10. Výpočtem v programu Statistica 10 jsme získali odlišné hodnoty, než při optimalizaci řešitelem. Odhady parametrů mají tyto hodnoty: $k = 241,56837$, $a = 71,00484$, $b = 0,50890$. S pravděpodobností 95% bude hodnota parametru modelu κ v intervalu $\langle 226,473; 256,663 \rangle$, hodnota parametru α v intervalu $\langle 22,37; 119,64 \rangle$ a hodnota parametru β $\langle 0,446; 0,572 \rangle$.

Tabulka 10: Statistica 10 – výstup č. 4

Model je: Prom2=(k/(1+a*b^Prom1)) (Tabulka4)						
Záv.prom.:Prom2						
Hladina spolehlivosti:95.0% (alfa =0.050)						
	Odhad	Standard chyba	t-hodn. sv = 9	p-hodn.	Dol. sp. Mez	Hor. sp. Mez
k	241,5684	6,67269	36,2025	0,00000	226,4737	256,6637
a	71,0048	21,4980	3,3028	0,00918	22,3729	119,6368
b	0,5089	0,0278	18,2816	0,00000	0,4459	0,5719

Zdroj: výstup z programu Statistica 10

Na následujícím grafu je zobrazena logistická křivka, která prolíná naměřené hodnoty y_i .



Graf 12: Statistica 10 - výstup č. 5

Zdroj: výstup z programu Statistica 10

$$\text{Získaná logistická rovnice má tvar: } y = \frac{241,568}{1 + 71,0048 \cdot 0,508897^x}.$$

Dalším výstupem, který je možné získat při výpočtu v programu Statistica je podíl rozptylu vysvětlený modelem neboli $r^2 = 0,994$ a $r = 0,997$.

V poslední části ilustrativního příkladu č. 2 bude pomocí interpolačních kritérií porovnán výpočet v programu Statistica 10 a výpočet dosažený optimalizací řešitelem. V následující tabulce budeme i nadále používat zkratky ŘEŠ pro řešitele a STA pro program Statistica 10.

Tabulka 11: Interpolační kritéria č. 4 - ilustrativní příklad č. 2

X	Y	Metody odhadu		MSE		MAE	
		Y(ŘEŠ)	Y(STA)	(ŘEŠ)	(STA)	(ŘEŠ)	(STA)
1	3,26	6,51	6,51	10,51	10,51	3,24	3,24
2	7,75	12,46	12,46	22,17	22,17	4,71	4,71
3	17,61	23,32	23,32	32,64	32,64	5,71	5,71
4	36,14	41,92	41,92	33,50	33,49	5,79	5,79
5	73,72	70,56	70,56	9,99	9,99	3,16	3,16
6	120,00	108,17	108,16	140,04	140,08	11,83	11,84
7	150,67	148,42	148,42	5,09	5,10	2,26	2,26
8	171,20	183,09	183,09	141,37	141,33	11,89	11,89
9	199,52	207,79	207,79	68,42	68,39	8,27	8,27
10	230,50	223,11	223,11	54,56	54,58	7,39	7,39
11	237,69	231,81	231,81	34,56	34,57	5,88	5,88
12	234,40	236,50	236,50	4,42	4,42	2,10	2,10
Suma				557,2617	557,2617	72,2307	72,2325
MSE / MAE				46,4385	46,4385	6,0192	6,0194
MIN MSE / MAE				46,4385	46,4385	6,0192	

Zdroj: vlastní zpracování v MS Excel

V tomto případě bylo proložení dat pomocí optimalizace řešitelem vhodnější než výpočet v programu Statistica 10. Ovšem rozdíly mezi interpolačními kritérii jsou pouze v řádech tisícín, a to konkrétně u kritéria MAE. Tato skutečnost jenom potvrzuje schopnost optimalizace v MS Excel co nejpřesněji proložit dané hodnoty růstovou křivkou. Hodnoty MSE a MAE jsou uvedeny v tabulce č. 11.

6.2.2 Predikce budoucích hodnot pomocí logistické rovnice

Pro odhad budoucích prodejů použijeme rovnici z výpočtu v programu Statistica 10 ve

tvaru: $Y = \frac{241,5687}{1 + 71,00535 \cdot 0,508897^x}$, kdy budeme odhadovat budoucí hodnoty prodejů

mobilních telefonů v Rusku pro roky 2012 a 2013.

Do rovnic dosadíme požadované x_i , v našem případě hodnoty 13 a 14. Predikce budoucích hodnot je následující:

$$y_{13} = \frac{241,5687}{1 + 71,00535 \cdot 0,508897^{13}} = 238,96$$

$$y_{14} = \frac{241,5687}{1 + 71,00535 \cdot 0,508897^{14}} = 240,24$$

Můžeme tedy říci, že na základě predikční rovnice, získané výpočtem v MS Excel a následnou optimalizací řešitelem, je odhad prodejů mobilních telefonů na území Ruska 238 990 000 kusů pro rok 2012 a 240 240 000 kusů pro rok 2013.

6.2.3 Extrapolace dat u logistické regrese

Podrobný popis této metody je uveden v kapitole 3.3.2. Pro náš příklad vynecháme poslední dvě pozorování a získáme následující náhodný výběr:

Tabulka 12: Extrapolace dat - náhodný výběr č. 2

x_i	1	2	3	4	5	6	7	8	9	10
y_i	3,263	7,75	17,61	36,14	73,72	120	150,7	171,2	199,5	230,5

Zdroj: [13]

Výpočtem parametrů predikční rovnice provedeným v programu Statistica 10 bylo dosaženo těchto hodnot: $a = 76,88$, $b = 0,4966$ a $k = 235,14$. Predikční rovnice má tedy tvar

$$Y = \frac{235,14}{1 + 76,88 \cdot 0,4966^x}$$

Odhad budoucích hodnot je následující:

$$y_{11} = \frac{235,14}{1 + 76,88 \cdot 0,4966^{11}} = 227,23$$

$$y_{12} = \frac{235,14}{1 + 76,88 \cdot 0,4966^{12}} = 231,15$$

Pro zhodnocení simulace použijeme opět Theilův koeficient, kdy po dosazení získáme:

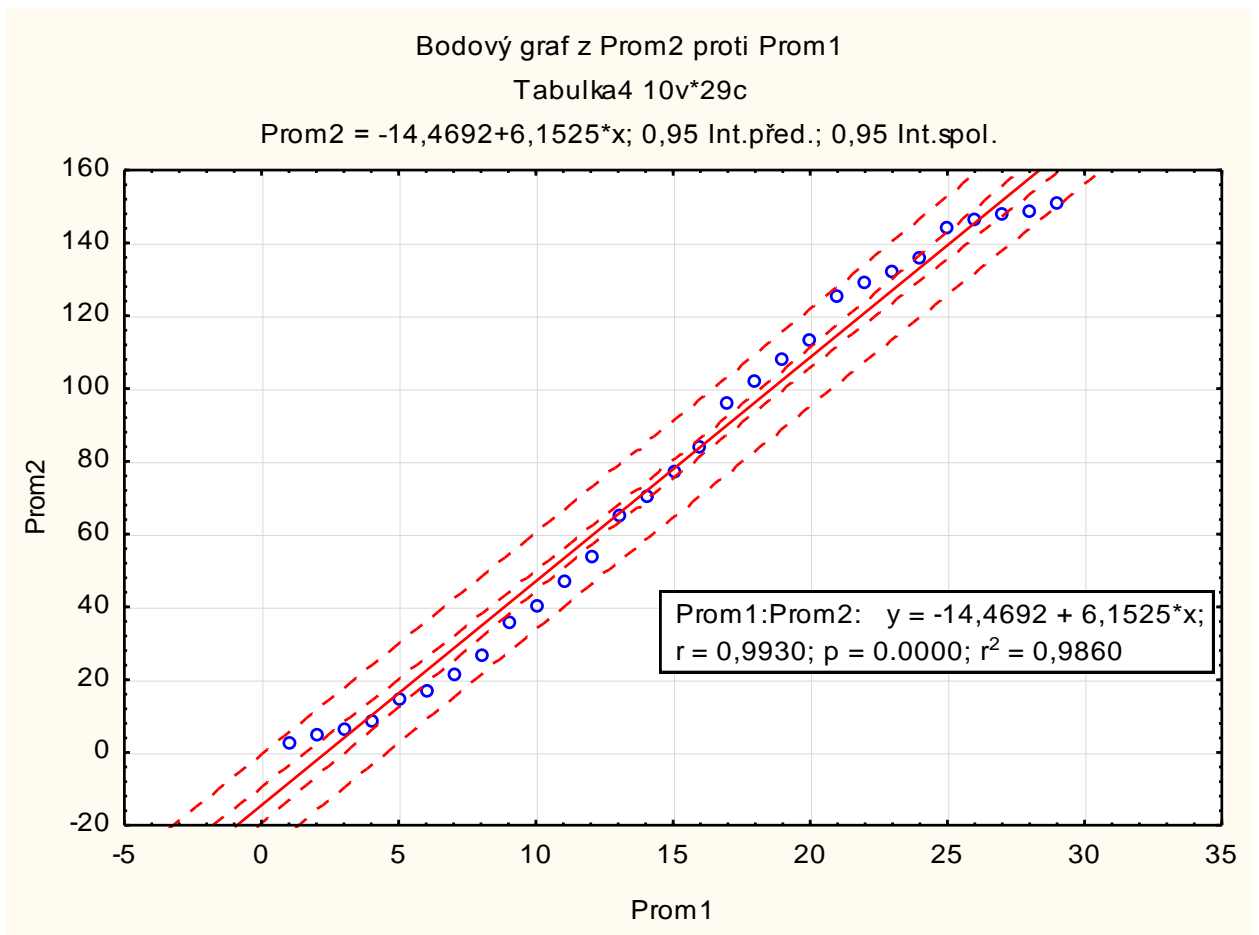
$$T_H^2 = \frac{(237,7 - 227,23)^2 + (234,4 - 231,15)^2}{237,7^2 + 234,4^2} = 0,00108$$

Na základě výsledné hodnoty Theilova koeficientu můžeme říci, že simulace i v druhém příkladu dosáhla dobrých výsledků. V absolutních hodnotách došlo u 11. období k odchylce 10,5 a u 12. období k odchylce 3,3.

6.3 Porovnání růstového a jednoduchého lineárního modelu

V této kapitole bude na ilustrativním příkladu č. 1 ukázána vhodnost využití růstového modelu, konkrétně Gompertzovy regrese oproti jednoduchému lineárnímu modelu. Na náhodný výběr obsahující čtvrtletní prodeje herní konzole NDS aplikujeme jednoduchou lineární regresi, přičemž cílem této části bude dokázat, že i přes velmi dobré výsledky, kterých bude aplikace lineárního modelu dosahovat, je vhodnější využít růstový model.

Připomeňme model jednoduché lineární regrese: $Y_i = \alpha + \beta X_i + \varepsilon_i$. Pro odhad parametrů α a β a proložení dat lineární přímkou použijeme program Statistica 10.



Graf 13: Jednoduchý lineární model

Zdroj: výstup z programu Statistica 10

Index determinace

Výstup zobrazený na grafu č. 11 získáme příkazem *Grafy – Bodový graf*. Predikční rovnice lineárního modelu má tvar $y = -14,47 + 6,15x$. Dále jsou na výstupu zobrazeny predikční pás a pás spolehlivosti. Širší z nich je pás predikční, který při zvolené hodnotě $\alpha = 0,05$ vymezuje oblast pro další naměřené hodnoty ze stejného souboru Y . Užší pás je pás spolehlivosti, který vymezuje na grafu oblast, ve které se s 95% pravděpodobností bude pohybovat nově vypočtená predikční funkce, pokud dojde k doplnění náhodného výběru o nově získané hodnoty. Dalším důležitým výstupem pro následné porovnání je hodnota $r^2 = 0,986$. Pokud porovnáme index determinace, který jsme získali při aplikaci Gompertzovy regrese, s indexem determinace získaným aplikací jednoduchou lineární regrese, zjistíme, že hodnota podílu rozptylu vysvětleného modelem je u Gompertzovy regrese vyšší než u lineární regrese. V předchozím textu jsme zjistili, že hodnota r^2 je rovna 0,9987. Na základě vyšší hodnoty indexu determinace, můžeme tvrdit, že Gompertzova křivka lépe přiléhá na dané

hodnoty, a tím lépe modeluje daný příklad, v našem případě čtvrtletní prodeje herní konzole NDS.

Extrapolace dat

Obdobným způsobem jako v předešlých výpočtech i zde použijeme simulaci extrapolací a následné vyhodnocení Theilovým koeficientem. V případě jednoduché lineární regrese máme možnost vyhodnotit vhodnost predikční funkce i díky výše uvedenému predikčnímu pásu. Vynecháme poslední dvě pozorování, čímž získáme soubor o 27 hodnotách. Vypočtená predikční funkce a predikční pás jsou zobrazeny na následujícím grafu.

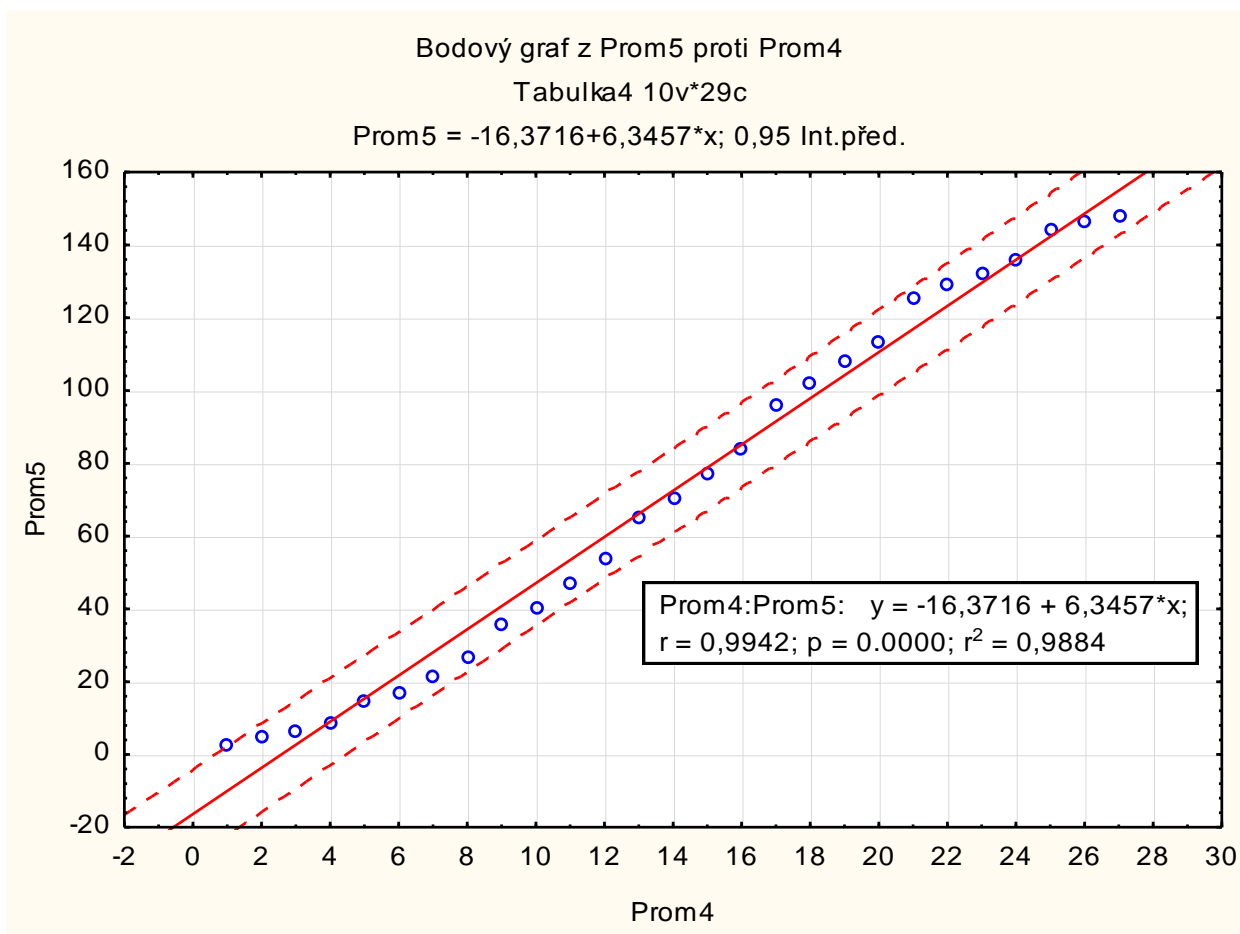
Získaná rovnice má tvar $y = -16,37 + 6,36x$. Dále můžeme vidět, že i v tomto případě po vynechání dvou měření je hodnota indexu determinace blízká jedné, kdy $r^2 = 0,9884$.

V programu Statistica 10 jsme jako výstup získali pouze absolutní hodnoty dolní i horní hranice predikčního pásu. Hranice pásu jsou stejně jako predikční rovnice jednoduchou lineární funkcí. Pomocným výpočtem v MS Excel získáme rovnici dolní i horní meze. Kdy použijeme označení y_{DM} pro dolní mez a y_{HM} pro horní mez. Rovnice mají následující tvar: $y_{DM} = -28,2297 + 6,3457x$; $y_{HM} = -4,5134 + 6,3457x$.

Nejprve provedeme zhodnocení simulace pomocí Theilova koeficientu. Po dosazení do predikční rovnice získáme hodnoty pro $y_{28} = 161,31$ a $y_{29} = 167,65$. Po dosazení do Theilova koeficientu získáme:

$$T_H^2 = \frac{(149 - 161,31)^2 + (151,06 - 167,65)^2}{149^2 - 161,31^2} = 0,0085$$

Výsledná hodnota koeficientu je vyšší než u simulace při využití růstového modelu kdy $T_H^2 = 0,00147$. Absolutní odchylky mezi predikovanými a vynechanými skutečnými hodnotami jsou 12,31 pro y_{28} a 16,59 pro y_{29} .



Graf 14: Simulace u jednoduché lineární regrese

Zdroj: výstup z programu Statistica 10

Pro použití predikčního pásu, jako jednoho z nástrojů vyhodnocení simulace lineární regrese, dosadíme do rovnic pro dolní mez a horní mez pásu a získáme intervaly pro $y_{28} = \langle 149,45; 173,17 \rangle$ a pro $y_{29} = \langle 155,8; 179,51 \rangle$, tyto intervaly nám při zvolené hodnotě $\alpha = 0,05$ určují interval, ve kterém se bude dané y_i nacházet.

Skutečné naměřené $y_{28} = 149,0$ nespadá do predikčního intervalu pro y_{28} a skutečné $y_{29} = 151,06$ se také nachází mimo daný interval. To je dalším důkazem toho, že je využití jednoduché lineární regrese nevhodné.

Závěrem této kapitoly můžeme tedy říci, že pokud data, které se rozhodneme modelovat jednou z dostupných metod, odpovídají svým charakterem růstovému modelu, je vhodnější použít jeden z růstových modelů a to i přes svůj náročnější výpočet. V našem případě bylo použitím jednoduché lineární regrese dosaženo dobrých interpolačních výsledků, ovšem následnou extrapolací bylo dosaženo nepřesných odhadů. Důvodem může být trend, který je typický pro ať už pro Gompertzovu nebo logistickou regresi. Po určitém časovém úseku, kdy

křivka roste určitým tempem, dojde do stádia, které je v životním cyklu výrobku označováno jako zralost popřípadě stagnace a dochází k výraznému zpomalení růstu. Pokud ovšem daný příklad modelujeme jednoduchou lineární regresí je tempo růstu konstantní, kdy je určeno směrnici křivky, neboli koeficientem β .

7 VYHODNOCENÍ JEDNOTLIVÝCH METOD

V předchozích kapitolách byl proveden výpočet a odhad parametrů Gompertzovy a logistické regrese celkem pěti možnými způsoby. Jednalo se o metodu dílčích průměrů, metodu částečných součtů, metodu vybraných bodů, výpočet řešitelem, výpočet programem Statistica 10 a u ilustrativního příkladu č. 1 byly naměřené hodnoty proloženy jednoduchou lineární regresí, zde se tedy jednalo o metodu nejmenších čtverců. Cílem této kapitoly je shrnout předešlé výpočty a výsledky a určit, která metoda odhadování parametrů je, jak pro Gompertzovu tak i pro logistickou křivku, nejvhodnější. Jinak řečeno, pomocí které metody, bylo dosaženo nejlepších výsledků.

7.1 Přehled výsledků Gompertzovy regrese

V prvním ilustrativním příkladu byla jako náhodná veličina vybrána čtvrtletní prodej herní konzole Nintendo DS. V následující tabulce je uvedeno shrnutí kritérií, na základě kterých je možno určit nejvhodnější metodu.

Tabulka 13: Vyhodnocení Gompertzovy regrese

Metoda	Funkce	Interpolace		Extrapolace	R ²
		MSE	MAE	Theilův koef.	
MDP	$y = 233,74 \cdot 0,0108^{0,912^x}$	35,70	3,94	-	0,988
MČS	$y = 188,28 \cdot 0,0235^{0,894^x}$	132,44	10,89	-	0,955
MVB	$y = 223,12 \cdot 0,206^{0,905^x}$	264,06	15,02	-	0,930
ŘEŠ	$y = 180,337 \cdot 0,00739^{0,88749^x}$	3,47	1,47	-	0,999
STA	$y = 180,337 \cdot 0,00739^{0,88749^x}$	3,47	1,47	0,0015	0,999
LIN*	$y = -14,47 + 6,15x$	37,30	5,20	0,0085	0,986

* jednoduchá lineární regrese

Zdroj: vlastní zpracování v MS Excel

Z tabulky je patrné, že nejlepších výsledků bylo dosaženo výpočtem v programu Statistica 10 (STA) a optimalizací řešitelem (ŘEŠ). Ovšem Theilův koeficient, jehož výpočet je podmíněn extrapolací, byl vypočítán pouze pomocí programu Statistica 10.

Výpočet indexu determinace pro MDP, MČS a MVB byl proveden v MS Excel na základě následujících vzorců [49]:

$$SSR = \sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2 \quad SSE = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad SST = SSR + SSE \quad R^2 = \frac{SSR}{SST}$$

SSR – vysvětlený součet čtverců odchylek

SSE – reziduální součet čtverců

SST – celkový součet čtverců

R² – index determinace

Nejméně kvalitních výsledků bylo dosaženo odhadem parametrů modelu pomocí metody vybraných bodů i přesto, že hodnota indexu determinace je 0,93.

Pro využití v praxi bychom za predikční rovnici, jakožto důležitý nástroj při manažerském rozhodování, zvolili rovnici dosaženou pomocí programu Statistica 10 nebo pomocí řešitele. Díky odhadu budoucích objemů prodeje má management určitou výhodu, jelikož má možnost přizpůsobit své rozhodování a řízení především co se týče výroby a distribuce právě predikovanému množství. Pokud je na základě predikční rovnice odhadován prodej přibližně 157 milionů kusů pro příští období, firma přizpůsobí objem výroby a distribuci, takovým způsobem, aby pro příští období bylo dané množství k dispozici.

7.2 Přehled výsledků logistické regrese

Logistickou regresí byl modelován soubor hodnot, které reprezentují roční prodeje mobilních telefonů v Rusku. Stejně jako u předchozí regrese i zde je uvedena tabulka, která uvádí jednotlivá hodnotící kritéria.

Tabulka 14: Vyhodnocení logistické regrese

Metoda	Funkce	Interpolace		Extrapolace	R ²
		MSE	MAE	Theilův koef.	
MDP	$y = \frac{145,558}{1 + 115,88 \cdot 0,380^x}$	2406,5700	33,5300	-	0,633
MČS	$y = \frac{232,024}{1 + 160,27 \cdot 0,429^x}$	94,1500	6,6700	-	0,988
MVB	$y = \frac{241,007}{1 + 171,5 \cdot 0,425^x}$	168,3800	7,2100	-	0,981
ŘEŠ	$y = \frac{241,5687}{1 + 71,00535 \cdot 0,508897^x}$	46,4385	6,0192	-	0,994
STA	$y = \frac{241,568}{1 + 71,0048 \cdot 0,508897^x}$	46,4385	6,0194	0,0011	0,994

Zdroj: vlastní zpracování v MS Excel

Pro výpočet jednotlivých indexů determinace byly použity stejné vzorce jako v předchozím příkladu. V tomto případě bylo na základě interpolačních kritérií dosaženo nejlepších výsledků pomocí řešitele. Ovšem i zde jsou výsledky dosažené řešitelem a programem Statistica 10 velmi podobné. Nebyla zde využita lineární regrese, jelikož z pouhého grafického zobrazení je viditelný růstový trend připomínající s-křivku, který je typický pro nelineární regrese. Nejhorších výsledků bylo v ilustrativním příkladu č. 2 dosaženo metodou dílčích průměrů, kdy hodnota indexu determinace je rovna 0,633.

Jelikož se jedná časovou řadu vystihující celkové prodeje všech druhů mobilních telefonů, využití této predikční funkce je odlišné než v předchozím případě. Odhad budoucích prodejů mobilních telefonů může využít potenciální nově vstupující prodejce nebo stávající prodejci. Pro rok 2012 je odhad přibližně 239 milionů prodaných kusů, kdy je určitý prodejce v daném regionu schopen odhadnout část z tohoto objemu, která případně právě jemu, pokud je mu známa informace o rozdělení trhu.

8 ZÁVĚR

Cílem této práce bylo vymežit logistickou a Gompertzovu regresi a to jak z hlediska teoretického a matematického, tak i z hlediska jejich využití v ekonomické praxi, kde růstové modely nacházejí stále větší uplatnění. Tento fakt souvisí se současným trendem pronikání ostatních vědních disciplín, jako jsou statistické a matematické metody do oblasti ekonomie. V první kapitole byly oba modely definovány a podrobněji popsány ze statistického hlediska. V další části práce byl uveden možný význam jednotlivých parametrů i z ekonomického hlediska. Využití těchto modelů spočívá v aplikaci daného modelu na vybrané časové řady. Pokud zkoumáme ekonomické využití, tak časovou řadou bývají zpravidla objemy prodeje vybraného výrobku za určité období na určitém území. Aplikace zvolené regrese spočívá v odhadu parametrů modelu a následném výběru nejvhodnějšího z nich pomocí interpolačních a extrapolacních kritérií. Tento postup byl podrobně popsán v kapitole 3.

Dalším cílem bylo tedy určit metodu odhadu, pomocí které je dosaženo nejlepších výsledků a to jak u Gompertzovy regrese, tak i u logistické regrese. Tomuto se věnuje praktická část diplomové práce, kdy jsou u obou modelů odhadnuty parametry pomocí metody dílčích průměrů, metody vybraných bodů, metody částečných součtů, řešitelem a iterační Levenberg-Marquardtovou metodou, kterou využívá program Statistica 10. Nejlepších výsledků u Gompertzovy regrese bylo na daném příkladu č. 1 dosaženo shodně pomocí řešitele i výpočtem v programu Statistica 10. U logistické regrese v příkladu č. 2 byl nejkvalitnější odhad parametrů proveden pomocí řešitele. Vyhodnocení bylo provedeno na základě indexů determinace a interpolačních kritérií jako jsou střední absolutní chyba odhadu (MAE) a střední čtvercová chyba odhadu (MSE). Pro tento účel jsou v závěrečné kapitole pro oba modely a jednotlivé metody odhadů použity přehledné tabulky, které srovnávají kritéria mezi jednotlivými metodami.

Hlavní přínos růstových modelů pro ekonomickou oblast spočívá v predikci budoucích hodnot, v našem případě jsme uvažovali odhady budoucích objemů prodeje u daných výrobců. Pokud tedy máme k dispozici výstup v podobě predikční rovnice, tak jsme schopni odhadnout budoucí prodeje. Tento odhad může znamenat například pro danou firmu jistou výhodu v manažerském rozhodování a v budoucích krocích firmy. Pokud je firma schopna na základě statistické analýzy předpovědět budoucí objemy prodeje svého výrobku dostává se i do určité konkurenční výhody, jelikož je oproti konkurenci schopna přizpůsobit dané procesy, spojené s výrobou a prodejem, právě odhadovaným číslům.

Veškeré cíle stanovené v úvodu práce byly splněny.

Použitá literatura

- [1] About Solver. In: Office.com: Excel . 2010. Dostupné z: <http://office.microsoft.com/en-us/excel-help/about-solver-HP005198368.aspx>
- [2] BERKOWITZ, Eric N., Roger A. KERIN a William RUDELIUS. Marketing. second edition. Homewood: Irwin, 1989, 752 s. ISBN 0-256-07513-1.
- [3] Call of Duty: Modern Warfare 3 - Xbox 360 - VGChartz. VGChartz. Dostupné z: <http://www.vgchartz.com/game/44606/call-of-duty-modern-warfare-3/Global/>
- [4] CRAVENS, David W., Gerald E. HILLS a Robert B. WOODRUFF. Marketing Management. Homewood: Irwin, 1987, 714 s. ISBN 0-256-05547-5.
- [5] GAVRILOVA, Natalia S. a Leonid A. GAVRILOV. Stárnutí a dlouhověkost: Zákony a prognózy úmrtnosti pro stárnoucí populace. Demografie . 2011, s. 20 . Dostupné z: <http://longevity-science.org/pdf/Demografie-2011.pdf>
- [6] GOMPERTZ, Benjamin. On the Nature of the Function Expressive of the Law of Human mortality, and on a New Mode of Determining the Value of life Contingencies. Philosophical Transactions of the Royal Society of London. 1825, vol. 115, s. 70.
- [7] HEBÁK, Petr, Jiří HUSTOPECKÝ a Ivana MALÁ. Vícerozměrné statistické metody. Vyd. 1. Praha: Informatorium, 2005, 239 s. ISBN 80-7333-036-9.
- [8] HINDLS, Richard, Jara KAŇOKOVÁ a Ilja NOVÁK. Metody statistické analýzy pro ekonomy. 1. vyd. Praha: Management Press, 1997, 249 s. ISBN 80-85943-44-1.
- [9] KINGSLAND, Sharon. THE JOHN'S HOPKINS UNIVERSITY. The refractory model: The logistic curve and the history of population ecology . Volume 57. Chicago: The University of Chicago Press, 1982 . Dostupné z: <http://www.jstor.org/discover/10.2307/2825134?uid=3737856&uid=2&uid=4&sid=56037070043>
- [10] KUBANOVÁ, Jana., Linda, Bohdan.: Porovnání vybraných metod odhadu parametrů Gompertzovy křivky. In: Sborník VI. mezinárodní konference“ Kvantitativne metody v ekonómii a podnikání“ FHI EU Bratislava 1999.
- [11] KUBANOVÁ, Jana. Statistické metody pro ekonomickou a technickou praxi. Vyd. 3., dopl. Bratislava: Statis, 2008, 247 s. ISBN 978-80-85659-47-4.

- [12] LIMPOUCH, Jiří. Levenberg-Marquardtova metoda. In: . s. 1, 2000-04-18 . Dostupné z: <http://www-troja.fjfi.cvut.cz/~limpouch/numet/extrem/node12.html>
- [13] Market Data: Asymco. ASYMCO . Horace Dediu. 2012. Dostupné z: <http://www.asymco.com/hire-me/vendor-bubbles/>
- [14] MATYS, Jaroslav. Gompertzova rozvojová křivka. Praha : Institut řízení, 1983. 88 s.
- [15] NINTENDO CO., Ltd. Consolidated financial highlights: 2004-12-31. 2007-04-26. Dostupné z: <http://www.nintendo.co.jp/ir/pdf/2005/050126e.pdf#page=6>
- [16] NINTENDO CO., Ltd. Consolidated financial highlights: 2005-03-31. 2007-04-26. Dostupné z: <http://www.nintendo.co.jp/ir/pdf/2005/050526e.pdf#page=39>
- [17] NINTENDO CO., Ltd. Consolidated financial highlights: 2005-06-30. 2007-01-25. Dostupné z: <http://www.nintendo.co.jp/ir/pdf/2005/050728e.pdf#page=6>
- [18] NINTENDO CO., Ltd. Consolidated financial highlights: 2005-09-30. 2007-04-26. Dostupné z: <http://www.nintendo.co.jp/ir/pdf/2005/051124e.pdf#page=25>
- [19] NINTENDO CO., Ltd. Consolidated financial highlights: 2005-12-31. 2007-04-26. Dostupné z: <http://www.nintendo.co.jp/ir/pdf/2006/060126e.pdf#page=6>
- [20] NINTENDO CO., Ltd. Consolidated financial highlights: 2006-03-31. 2007-04-26. Dostupné z: <http://www.nintendo.co.jp/ir/pdf/2006/060525e.pdf#page=30>
- [21] NINTENDO CO., Ltd. Consolidated financial highlights: 2006-06-30. 2007-01-25. Dostupné z: <http://www.nintendo.co.jp/ir/pdf/2006/060724e.pdf#page=8>
- [22] NINTENDO CO., Ltd. Consolidated financial highlights: 2006-09-30. 2007-01-25. Dostupné z: <http://www.nintendo.co.jp/ir/pdf/2006/061026e.pdf#page=28>
- [23] NINTENDO CO., Ltd. Consolidated financial highlights: 2006-12-31. 2007-01-25. Dostupné z: <http://www.nintendo.co.jp/ir/pdf/2007/070125e.pdf#page=6>
- [24] NINTENDO CO., Ltd. Consolidated financial highlights: 2007-03-31. 2007-04-26. Dostupné z: <http://www.nintendo.co.jp/ir/pdf/2007/070426e.pdf#page=21>
- [25] NINTENDO CO., Ltd. Consolidated financial highlights: 2007-06-30. 2007-07-25. Dostupné z: <http://www.nintendo.co.jp/ir/pdf/2007/070725e.pdf#page=8>
- [26] NINTENDO CO., Ltd. Consolidated financial highlights: 2007-09-30. 2007-07-25. Dostupné z: <http://www.nintendo.co.jp/ir/pdf/2007/071025e.pdf#page=22>

- [27] NINTENDO CO., Ltd. Consolidated financial highlights: 2007-12-31. 2007-01-24.
Dostupné z: <http://www.nintendo.co.jp/ir/pdf/2008/080124e.pdf#page=8>
- [28] NINTENDO CO., Ltd. Consolidated financial highlights: 2008-03-31. 2008-04-24.
Dostupné z: <http://www.nintendo.co.jp/ir/pdf/2008/080424e.pdf#page=22>
- [29] NINTENDO CO., Ltd. Consolidated financial highlights: 2008-06-30. 2008-07-30.
Dostupné z: <http://www.nintendo.co.jp/ir/pdf/2008/080730e.pdf#page=11>
- [30] NINTENDO CO., Ltd. Consolidated financial highlights: 2008-09-30. 2008-10-30.
Dostupné z: <http://www.nintendo.co.jp/ir/pdf/2008/081030e.pdf#page=11>
- [31] NINTENDO CO., Ltd. Consolidated financial highlights: 2008-12-31. 2009-01-29.
Dostupné z: <http://www.nintendo.co.jp/ir/pdf/2009/090129e.pdf#page=11>
- [32] NINTENDO CO., Ltd. Consolidated financial highlights: 2009-03-31. 2009-04-07.
Dostupné z: <http://www.nintendo.co.jp/ir/pdf/2009/090507e.pdf#page=23>
- [33] NINTENDO CO., Ltd. Consolidated financial highlights: 2009-06-30. 2009-07-30.
Dostupné z: <http://www.nintendo.co.jp/ir/pdf/2009/090730e.pdf#page=9>
- [34] NINTENDO CO., Ltd. Consolidated financial highlights: 2009-09-30. 2009-10-29.
Dostupné z: <http://www.nintendo.co.jp/ir/pdf/2009/091029e.pdf#page=9>
- [35] NINTENDO CO., Ltd. Consolidated financial highlights: 2009-12-31. 2010-01-28.
Dostupné z: <http://www.nintendo.co.jp/ir/pdf/2010/100128e.pdf#page=9>
- [36] NINTENDO CO., Ltd. Consolidated financial highlights: 2010-03-31. 2010-05-07.
Dostupné z: <http://www.nintendo.co.jp/ir/pdf/2010/100506e.pdf#page=23>
- [37] NINTENDO CO., Ltd. Consolidated financial highlights: 2010-06-30. 2010-08-05.
Dostupné z: <http://www.nintendo.co.jp/ir/pdf/2010/100729e.pdf#page=9>
- [38] NINTENDO CO., Ltd. Consolidated financial highlights: 2010-09-30. 2010-10-28.
Dostupné z: <http://www.nintendo.co.jp/ir/pdf/2010/101028e.pdf#page=9>
- [39] NINTENDO CO., Ltd. Consolidated financial highlights: 2010-12-31. 2011-06-05.
Dostupné z: <http://www.nintendo.co.jp/ir/pdf/2011/110127e.pdf#page=9>
- [40] NINTENDO CO., Ltd. Consolidated financial highlights: 2011-03-31. 2010-06-05.
Dostupné z: <http://www.nintendo.co.jp/ir/pdf/2011/110425e.pdf#page=16>

- [41] NINTENDO CO., Ltd. Consolidated financial highlights: 2011-06-30. 201-01-26.
Dostupné z:
http://www.nintendo.co.jp/ir/library/historical_data/pdf/consolidated_sales_e1106.pdf
- [42] NINTENDO CO., Ltd. Consolidated financial highlights: 2011-09-30. 2011-10-27.
Dostupné z:
http://www.nintendo.co.jp/ir/library/historical_data/pdf/consolidated_sales_e1109.pdf
- [43] NINTENDO CO., Ltd. Consolidated financial highlights: 2011-12-31. 2012-01-26.
Dostupné z:
http://www.nintendo.co.jp/ir/library/historical_data/pdf/consolidated_sales_e1112.pdf
- [44] NUNES, Paul. Jumping the S-curve: How to beat the growth cycle, get on top and stay there. Boston : Harvard Business School Publishig, 2011. 270 s. ISBN 978-0-471-47135-6
- [45] PlayStation 3 Worldwide Software Unit Sales: CORPORATE INFORMATION. SONY COMPUTER ENTERTAINMENT. Sony Computer Entertainment Inc. Dostupné z:
http://www.scei.co.jp/corporate/data/bizdataps3soft_sale_e.html
- [46] RUBLÍKOVÁ, Eva. Analýza časových radov. 1. vyd. Bratislava: Iura Edition, 2007, 207 s. ISBN 978-80-8078-139-2.
- [47] SEBER, George A. F. Nonlinear regression. New Jersey : John Wiley and Sons, Inc. 2003. 792 s. ISBN 978-0-471-47135-6
- [48] VYSEKALOVÁ, Jitka. Chování zákazníka: jak odkrýt tajemství "černé skříňky". 1. vyd. Praha : Grada, 2011. 356 s. ISBN 978-80-247-3528-3
- [49] WALPOLE, Ronald E. et al Probability and statistics for engineers and scientist. 7th ed. New Jersey: Prentice-Hall, Inc., 2002, 730 s. ISBN 0-13-041529-4.