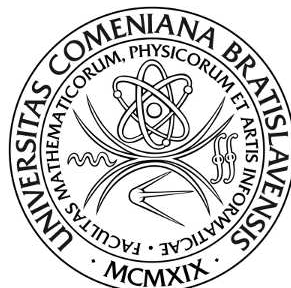


UNIVERZITA KOMENSKÉHO V BRATISLAVE  
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

NÁVRHY V KLINICKÝCH SKÚŠKACH  
S OHRANIČENIAMI

BAKALÁRSKA PRÁCA

UNIVERZITA KOMENSKÉHO V BRATISLAVE  
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY



NÁVRHY V KLINICKÝCH SKÚŠKACH  
S OHRANIČENIAMÍ

BAKALÁRSKA PRÁCA

Študijný program: Poistná matematika  
Študijný odbor: 6211 Štatistika  
Školiace pracovisko: Katedra aplikovanej matematiky a štatistiky  
Vedúci práce: Mgr. Lenka Filová, PhD.



Univerzita Komenského v Bratislave  
Fakulta matematiky, fyziky a informatiky

---

## ZADANIE ZÁVEREČNEJ PRÁCE

**Meno a priezvisko študenta:** Patrícia Blažová  
**Študijný program:** poistná matematika (Jednoodborové štúdium, bakalársky I. st., denná forma)  
**Študijný odbor:** štatistika  
**Typ záverečnej práce:** bakalárska  
**Jazyk záverečnej práce:** slovenský  
**Sekundárny jazyk:** anglický

**Názov:** Návrhy v klinických skúškach s ohraničeniami  
*Clinical trial designs with constraints*

**Anotácia:** Návrhy experimentov pre odhadovanie vzťahu medzi dávkou a odozvou pacienta mávajú často ohraničenia vynútené klinickou praxou. Jedným z takýchto ohraničení je to, aby pravdepodobnosť podania neúčinnnej alebo toxickej dávky nepresiahla nejakú hodnotu. V práci budeme hľadať optimálne návrhy v situácii, keď sú tieto ohraničenia lineárne a skúmať ich efektivitu vzhľadom na návrhy bez ohraničení.

**Vedúci:** Mgr. Lenka Filová, PhD.  
**Katedra:** FMFI.KAMŠ - Katedra aplikovanej matematiky a štatistiky  
**Vedúci katedry:** prof. RNDr. Marek Fila, DrSc.  
**Dátum zadania:** 09.10.2019

**Dátum schválenia:** 16.10.2019

doc. RNDr. Katarína Janková, CSc.  
garant študijného programu

.....  
študent

.....  
vedúci práce

**Podakovanie** Touto cestou by som sa chcela poďakovať školiteľke mojej bakalárskej práce Mgr. Lenke Filovej, PhD. za cenné rady a pripomienky, no najmä za ochotu, s akou pristupovala k odbornému vedeniu a v neposlednom rade za čas, ktorý venovala mojej práci.

# Abstrakt

BLAŽOVÁ, Patrícia: Návrhy v klinických skúškach s ohraničeniami [Bakalárska práca], Univerzita Komenského v Bratislave, Fakulta matematiky, fyziky a informatiky, Katedra aplikovanej matematiky a štatistiky; vedúci práce: Mgr. Lenka Filová, PhD., Bratislava, 2020, 38 s.

V tejto práci sa venujeme navrhovaniu experimentov, ktoré sú súčasťou klinických štúdií, zaoberajúcich sa testovaním nového lieku predtým, než bude uvedený do praxe. Vzťah medzi dávkou a odozvou pacienta odhadujeme pomocou nelineárnych modelov, ktoré popisujú pravdepodobnosť podania toxickéj alebo neúčinnejšej dávky. Za predpokladu lineárnych ohraničení hľadáme exaktné optimálne návrhy, ktoré popri daných obmedzeniach predstavujú najlepšie riešenie, čo sa týka správneho podávania liečiva. Tieto návrhy nájdeme pomocou kritérií optimality a ich efektivitu skúmame vzhľadom na teoretické optimum, čiže návrhy bez ohraničení. Výpočty a výsledky experimentu, sú realizované v štatistickom softvéri R.

**Kľúčové slová:** optimálny návrh, lineárne ohraničenie, efektivita, kritérium optimality

# Abstract

BLAŽOVÁ, Patrícia: Clinical trial designs with constraints [Bachelor Thesis], Comenius University in Bratislava, Faculty of Mathematics, Physics and Informatics, Department of Applied Mathematics and Statistics; Supervisor: Mgr. Lenka Filová, PhD., Bratislava, 2020, 38 p.

In this thesis, we consider design of experiments, which are part of clinical trials dealing with testing a new drug before it will be used in practice. We estimate relationship between dose and patient response using nonlinear models, which interpret probability of giving toxic or inefficient dose. On assuming linear constraints, we search for exact optimal designs, that are the best solution of treatment dosage with respect to the constraints. We find these designs using optimality criteria and study their efficiency with respect to the theoretical optimum, i.e., designs without constraints. The final part of this thesis shows calculations and results of the experiment, which is implemented in statistical software R.

**Keywords:** optimal design, linear constraint, efficiency, optimality criterion

# Obsah

Úvod	7
<b>1 Klinické skúšky</b>	<b>8</b>
<b>2 Navrhovanie experimentu</b>	<b>10</b>
2.1 Lineárny model . . . . .	10
2.2 Návrh experimentu . . . . .	11
2.2.1 Aproximatívny a exaktný návrh . . . . .	11
2.2.2 Informačná matica . . . . .	12
2.2.3 Kritériá optimality . . . . .	12
2.2.4 Efektívnosť návrhu . . . . .	14
2.2.5 Návrh s ohraničením . . . . .	16
<b>3 Nelineárny model</b>	<b>19</b>
3.1 Logistický model . . . . .	20
3.1.1 Štvorparametrový logistický model . . . . .	22
3.2 EMAX model . . . . .	22
3.3 Log-lineárny model . . . . .	23
<b>4 Realizácia experimentu</b>	<b>24</b>
4.1 A- a D-optimálne návrhy pre dvojparametrový logistický model . . . . .	24
4.2 A-a D-optimálne návrhy pre štvorparametrový logistický model . . . . .	29
4.3 A- a D-optimálne návrhy pre log-lineárny model a EMAX model . . . . .	31
<b>Záver</b>	<b>35</b>
<b>Zoznam použitej literatúry</b>	<b>37</b>
<b>Prílohy</b>	

# Úvod

Vývoju nového lieku a jeho následnému zavedeniu do praxe predchádza dlhé obdobie skúmania a testovania v laboratóriách. Až po týchto úspešných testoch sa môže pristúpiť ku klinickým skúškam, ktoré však môžu vyžadovať isté ohraničenie, napr. sú limitované finančné prostriedky alebo počet dobrovoľníkov zúčastňujúcich sa testovania. Aby sa zabránilo podaniu dávky, ktorá je toxická alebo naopak aby sa zbytočne nepodávala neúčinná dávka, skúmame súvislosť medzi dávkou a odozvou pacienta pomocou rôznych modelov.

Hlavným účelom tejto práce je nájsť optimálne návrhy experimentov pre klinické skúšky testovania liekov za predpokladu lineárnych ohraničení a skúmať ich efektivitu vzhľadom na návrhy bez ohraničení.

Bakalárska práca je členená na 4 kapitoly. V prvej z nich sú zhrnuté jednotlivé fázy, z ktorých sa skladajú klinické štúdie a podľa ktorých sa postupuje pri zavádzaní nových medikamentov do praxe.

V druhej kapitole sú uvedené základné poznatky a vlastnosti lineárneho modelu a druhú časť tvoria pojmy ako aproximatívny a exaktný návrh, informačná matica, kritériá optimality, z ktorých sú predstavené kritériá A a D spolu s efektivitou a záver kapitoly predstavuje návrh s ohraničením.

Tretia kapitola sa zaoberá nelineárnymi modelmi, ktoré v tejto práci popisujú pravdepodobnosť toxicity a efektivity dávok. Nachádzajú sa v nej definície a súčasne vlastnosti log-lineárneho modelu, EMAX modelu a bližšie je priblížený logistický model, pri ktorom sú uvedené viaceré jeho typy.

Dôležitú časť bakalárskej práce tvorí záverečná kapitola, ktorá tvorí praktickú časť, teda realizáciu experimentu. Sú v nej uvedené výsledky optimálnych návrhov, ktoré za daných lineárnych ohraničení predstavujú správne dávkovanie liečiva a súčasne zabraňujú situácii, aby pacient dostal dávku, ktorá by mala vedľajšie účinky. Taktiež sú tu uvedené hodnoty efektivity týchto návrhov vzhľadom na aproximatívne návrhy bez ohraničení.



# 1 Klinické skúšky

Každý liek, pred tým než sa dostane k pacientovi, musí prejsť klinickými testami. Tým zase predchádzajú preklinické štúdie, ktorých úlohou je testovanie nového lieku z hľadiska bezpečnosti a ďalších možných rizík. Iba v prípade, že je zaručená bezpečnosť, sa môže pristúpiť ku klinickým skúškam, pri ktorých sa liek testuje na dobrovoľníkoch. Základné postupy a princípy klinických skúšok, ktoré sa zakladajú na tzv. „Správnej klinickej praxi“, sú rovnaké prakticky na celom svete. Klinické vyhodnocovanie liekov môžeme rozdeliť do štyroch fáz [4]:

**1.fáza** - ide o prvé podanie lieku zdravým dobrovoľníkom, ktorých býva niekoľko desiatok a ktorým sa za účasť v tejto fáze väčšinou platí. Cieľom je určiť maximálnu tolerovateľnú dávku, ktorú ľudské telo dokáže ešte prijať a získať informácie o znášanlivosti liečiva, ktoré slúžia ako podklad pre určenie dávky v nasledujúcej fáze.

**2.fáza** - ak sa v prvej fáze neobjavili žiadne závažné riziká, prechádza sa do druhej fázy, v ktorej sa skúmaný liek podáva pacientom trpiacim danou chorobou. Zisťujú sa nežiadúce účinky, vyhodnocujú sa predpokladané indikácie a sledujú sa zlepšenia zdravotného stavu pacientov. Počet zúčastnených pacientov sa pohybuje okolo stovky až niekoľko stoviek. Táto fáza trvá rok a pol, a ak sa potvrdí, že liek zlepšil liečbu daného ochorenia, prejde sa do tretej fázy testovania.

**3.fáza** - nazýva sa aj Rozšírená klinická štúdia a je to najnáročnejšia časť testovania. V tejto fáze pozorujeme terapeutickú účinnosť a bezpečnosť hodnoteného lieku, pričom ho porovnávame už s používanými liekmi, ktoré sú považované za najúčinnšie pre danú liečbu. Do tejto fázy je zahrnutých niekoľko stoviek až tisíc pacientov z viacerých krajín, aby sa získali štatisticky hodnotné údaje o bezpečnosti lieku. Na základe výsledkov všetkých fáz klinických skúšok, ktoré sa predložia kontrolným úradom (u nás Štátny ústav pre kontrolu liečiv), môže byť liek zaregistrovaný a následne používaný v klinickej praxi.

**4.fáza** - je to fáza overovania účinkov lieku v bežnej klinickej praxi za normálnych terapeutických podmienok v prvých rokoch od jeho registrácie. Sledujú sa nežiadúce účinky v rôznych vekových skupinách pacientov, interakcie s inými liekmi, prípadne sa vytvárajú nové dávky a kombinácie s liekmi, ktoré sú štandardne používané pri liečbe

rovnakého ochorenia.

V našej práci sa budeme venovať matematickým problémom vyskytujúcim sa v druhej fáze, konkrétne budeme sledovať vedľajšie nežiadúce účinky lieku. Zároveň sa budeme snažiť nastaviť optimálne dávkovanie daného liečiva s využitím viacerých modelov, ktoré sa používajú v klinických skúškach, tak aby sme predišli tomu, že pacientovi podáme dávku, ktorá by mala toxický účinok. Takisto budeme zisťovať účinnosť dávok pri daných ohraničeniach, napr. keď máme k dispozícii len obmedzený počet pacientov alebo keď máme limitovaný finančný rozpočet, čo býva najčastejší problém, keďže v súčasnosti je testovanie nového lieku veľmi nákladné.

## 2 Navrhovanie experimentu

Ako sme už spomínali, v klinických skúškach sa snažíme nájsť optimálny vzťah medzi dávkou testovaného lieku a odozvou pacienta. Na to sa však používa viacero modelov, v ktorých potrebujeme čo najpresnejšie odhadnúť neznáme parametre modelu. Kvalita týchto odhadov ale bude závisieť od návrhu experimentu, teda od toho, akým spôsobom sme namerali dáta (aké dávky sme podávali, akému množstvu pacientov a podobne). Celý postup si kvôli jednoduchosti najskôr ukážeme na lineárnom modeli.

### 2.1 Lineárny model

Lineárny regresný model môžeme definovať v maticovom tvare predpisom [11]:

$$y = F\theta + \varepsilon \quad (2.1)$$

pričom  $F$  je matica taká, že

$$F = \begin{pmatrix} f^T(x_1) \\ \vdots \\ f^T(x_N) \end{pmatrix}$$

alebo v tvare

$$y_k(x_k) = f^T(x_k)\theta + \varepsilon_k \quad k = 1, \dots, N \quad (2.2)$$

kde  $x = (x_1, \dots, x_N)^T$  je vektor nezávislých premenných, ktorý obsahuje body návrhu experimentu,  $y = (y_1, \dots, y_N)^T$  je vektor závislých premenných, ktorý obsahuje výsledky merania a  $\varepsilon \in \mathbb{R}^N$  je náhodný vektor (vektor chýb) taký, že

$$E(\varepsilon_k) = 0, \quad \text{var}(\varepsilon_k) = \sigma^2.$$

Vektor  $x \in \mathcal{X}$ , kde  $\mathcal{X}$  je kompaktná množina všetkých pokusov experimentu [1].

Naším cieľom je na základe realizácií merania odhadnúť vektor neznámych parametrov  $\theta$  metódou najmenších štvorcov, t.j. vypočítať:

$$\hat{\theta} = \arg \min_{\theta \in \mathbb{R}^m} \sum_{k=1}^N [y_k - f^T(x_k)\theta]^2 \quad (2.3)$$

Ak  $N > m$  a matica  $F^T F$  je regulárna, potom riešením rovnice (2.3) je:

$$\hat{\theta} = (F^T F)^{-1} F^T y, \quad (2.4)$$

pričom  $\hat{\theta}$  je nevychýlený odhad  $\theta$ , t.j.  $E(\hat{\theta}) = \theta$  a kovariančná matica pre  $\hat{\theta}$  má tvar  $\text{Cov}(\hat{\theta}) = \sigma^2(F^T F)^{-1}$ .

## 2.2 Návrh experimentu

Uvažujme množinu  $\mathcal{X} = \{x_1, \dots, x_n\}$ , ktorá bude predstavovať všetky možné body návrhu nášho experimentu. Rozsah tejto množiny môže byť veľmi veľký, aj niekoľko sto tisíc bodov. Našou úlohou je vybrať z tejto množiny niekoľko bodov (napr. 100), ktoré nám zaručia, že budeme vedieť dostatočne presne odhadnúť neznáme parametre. Takýto výber nazývame návrh experimentu. V každom vybranom bode návrhu môžeme vykonať viacero meraní, preto si musíme najskôr určiť s akým návrhom budeme pracovať. Nasledujúce poznatky sme čerpali z [14].

### 2.2.1 Aproximatívny a exaktný návrh

Aproximatívny návrh definujeme ako pravdepodobnostnú mieru  $\xi$  na množine  $\mathcal{X}$ . Tento návrh sa zvykne nazývať aj ako spojitý a označuje sa

$$\xi = \left\{ \begin{array}{ccc} x_1 & \dots & x_n \\ w_1 & \dots & w_n \end{array} \right\} \quad (2.5)$$

kde  $x_1, \dots, x_n$  predstavujú body návrhu a  $w_i$  sú váhy, ktoré označujú pomer celkových meraní uskutočnených v bodoch  $x_i$ , teda  $w_i = \xi(x_i)$ , pričom  $\sum_{i=1}^n w_i = 1, 0 \leq w_i \leq 1$ . Množinu všetkých prípustných návrhov na  $\mathcal{X}$  označíme  $\Xi$ .

Ak je návrhová miera realizovateľná v celých číslach, hovoríme o exaktnom návrhu experimentu, ktorý značíme

$$\xi_N = \left\{ \begin{array}{ccc} x_1 & \dots & x_n \\ m_1/N & \dots & m_n/N \end{array} \right\} \quad (2.6)$$

kde  $N$  je celkový počet meraní,  $m_i$  je počet meraní v bode  $x_i$ ,  $m_i \in \mathbb{N}$  a  $\sum_{i=1}^n m_i = N$ . Potom funkciu  $\xi$  v jednotlivých bodoch  $x_i$  vieme zapísať ako  $\xi(x_i) = m_i/N$ . V praktických situáciách pracujeme predovšetkým s exaktnými návrhmi, ale aproximatívne návrhy sa používajú z dôvodu menšej výpočtovej náročnosti pri ich hľadaní.

## 2.2.2 Informačná matica

Informačnú maticu typu  $m \times m$  pre návrh  $\xi$  definujeme nasledovne [14]:

$$M(\xi, \theta) = \sum_{i=1}^n M(x_i, \theta) w_i, \quad (2.7)$$

kde  $M(x_i, \theta)$  je informačná matica merania v bode  $x_i$ , ktorú môžeme zapísať v tvare  $M(x_i, \theta) = f(x_i, \theta) f^T(x_i, \theta)$ . Informačné matice  $M(\xi, \theta)$  sú pozitívne semidefinitné.

Ak uvažujeme lineárny regresný model z kapitoly 2.1, informačná matica návrhu  $\xi$  nezávisí od parametrov  $\theta$ , teda

$$M(\xi) = \sum_{x \in \mathcal{X}} f(x) f^T(x) \xi(x). \quad (2.8)$$

Cielom optimálneho návrhu je pri danom vektore, ktorý obsahuje body návrhu experimentu, maximalizovať presnosť odhadovaného vektora  $\theta$ , prípadne jeho funkcií alebo analogicky minimalizovať kovariančnú maticu  $\text{Cov}(\hat{\theta})$  odhadovaných parametrov. V nelineárnych modeloch, ktoré uvedieme neskôr, kovariančnú maticu nepoznáme, ale vieme ju aproximovať inverznou maticou k Fisherovej informačnej matici, teda môžeme písať

$$\text{Cov}(\hat{\theta}) = M^{-1}(\xi, \theta). \quad (2.9)$$

## 2.2.3 Kritériá optimality

V našej práci sa budeme snažiť pomocou informačnej matice maximalizovať množstvo informácií získaných z experimentu. Preto sa zameriame na maximalizáciu funkcií informačnej matice. Pri hľadaní optimálneho návrhu sa používajú rôzne funkcie. Tieto funkcie  $\Phi : \mathcal{S}^m \rightarrow \mathbb{R}$ , kde  $\mathcal{S}^m$  je množina pozitívne semidefinitných matíc typu  $m \times m$ , označujeme ako kritériá optimality. Návrh  $\xi^*$  je  $\Phi$ -optimálny, ak maximalizuje hodnotu kritéria, teda

$$\xi^* = \arg \max_{\xi \in \Xi} \Phi(M(\xi)). \quad (2.10)$$

Existuje viacero kritérií optimality, pomocou ktorých vieme získať optimálny návrh. My budeme pri našom experimente hľadať návrhy, ktoré spĺňajú kritériá D a A. Ďalšie kritériá optimality nájde čitateľ v [1],[3] a [14].

### D-optimalita:

D-optimálny návrh je taký návrh, ktorý maximalizuje determinant informačnej matice, t.j.

$$\xi_D = \arg \max_{\xi \in \Xi} \det M(\xi). \quad (2.11)$$

Kritérium D-optimality je definované v tvare:  $\Phi_D(M(\xi)) = (\det M(\xi))^{1/m}$ .

Toto kritérium má užitočné matematické vlastnosti. D-optimálny návrh je invariantný voči lineárnej transformácii a nezávisí na parametrizácii daného modelu. To znamená, že ak zmeníme konštrukčný interval, môžeme odvodiť D-optimálny návrh priamo z toho, ktorý bol skonštruovaný v pôvodnom intervale návrhu. Takúto vlastnosť nemusia spĺňať ostatné kritériá optimality, a to je jeden z dôvodov prečo sa kritérium D-optimality používa tak často. Nasledujúce tvrdenie ukazuje, akým spôsobom môžeme hľadať D-optimálne návrhy.

**Tvrdenie 1.** [12], [14] *Označme  $\xi_D$  D-optimálny návrh. Potom nasledujúce tri podmienky sú ekvivalentné:*

- $\xi_D = \arg \max_{\xi \in \Xi} \Phi(M(\xi))$ ,
- $\xi_D = \arg \min_{\xi \in \Xi} \max_{x \in \mathcal{X}} d(x, \xi)$ ,
- $\max_{x \in \mathcal{X}} d(x, \xi_D) = m$ .

*Funkciu  $d(x, \xi) = f^T(x)M^{-1}(\xi)f(x)$  nazývame variančná funkcia návrhu  $\xi$  pre  $x \in \mathcal{X}$ .*

Toto tvrdenie sa zvykne nazývať aj veta o ekvivalencii a je to jedno z najdôležitejších tvrdení v oblasti navrhovania optimálneho návrhu, pričom sa dá zovšeobecniť aj pre iné kritériá optimality.

### A-optimalita:

V prípade A-optimality sa snažíme dosiahnuť, aby celková disperzia odhadovaných parametrov, resp. priemerná disperzia, bola čo najmenšia. Preto treba minimalizovať stopu kovariančnej matice, čo môžeme zapísať ako

$$\xi_A = \arg \min_{\xi \in \Xi} \text{tr}[M^{-1}(\xi)], \quad (2.12)$$

a to je ekvivalentné s formou kritéria:  $\Phi_A(M(\xi)) = m/\text{tr}[M^{-1}(\xi)]$ .

Na rozdiel od D-optimality, nevýhodou tohto kritéria je, že ak preškalujeme pôvodný konštrukčný interval, môžeme dostať úplne iný optimálny návrh. V niektorých aplikáciách však táto nevýhoda nemusí byť taká dôležitá.

#### 2.2.4 Efektívnosť návrhu

Pri návrhu experimentu nás bude zaujímať do akej miery je náš návrh efektívny vzhľadom na optimálny návrh. Efektivita nám umožňuje určiť, na koľko percent je náš návrh “dobrý” v porovnaní s optimálnym návrhom.

**Definícia 1.** [14] *D-efektívnosť návrhu  $\xi$  je definovaná ako*

$$D_{eff}(\xi) = \left( \frac{\det M(\xi)}{\det M(\xi_D)} \right)^{\frac{1}{m}} \in [0, 1], \quad (2.13)$$

kde  $\xi_D$  je D-optimálny návrh a  $m$  je počet parametrov daného modelu.

**Definícia 2.** [5] *A-efektívnosť návrhu  $\xi$  definujeme ako*

$$A_{eff}(\xi) = \frac{\text{tr} M^{-1}(\xi_A)}{\text{tr} M^{-1}(\xi)} \in [0, 1], \quad (2.14)$$

kde  $\xi_A$  predstavuje A-optimálny návrh.

**Príklad 1.** Uvažujme jednoduchý lineárny model s dvomi parametrami tvaru  $y_i = \theta_0 + \theta_1 x_i + \varepsilon_i$  pre  $i = 1, \dots, N$ , pre ktorý matica plánu má tvar

$$F = \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}.$$

Ďalej uvažujme množinu  $\mathcal{X} = \{0, 0.25, 0.5, 0.75, 1\}$ , ktorá predstavuje dávky lieku, pričom pacientom pridáme rovnomerne dávky

$$\xi = \begin{Bmatrix} 0 & 0.25 & 0.5 & 0.75 & 1 \\ 1/5 & 1/5 & 1/5 & 1/5 & 1/5 \end{Bmatrix}.$$

Informačná matica pre tento návrh má tvar

$$M(\xi) = \sum_{i=1}^5 (1 \ x_i)^T (1 \ x_i) \xi(x_i) = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 0.375 \end{pmatrix}.$$

Maximalizujeme determinant informačnej matice a následne dostávame D-optimálny návrh

$$\xi_D = \begin{Bmatrix} 0 & 0.25 & 0.5 & 0.75 & 1 \\ 1/2 & 0 & 0 & 0 & 1/2 \end{Bmatrix}.$$

Hodnoty variančnej funkcie štandardného návrhu sú

$$d(x, \xi) = (3.0, 1.5, 1.0, 1.5, 3.0)^T.$$

Pre hodnoty variančnej funkcie optimálneho návrhu dostaneme

$$d(x, \xi_D) = (2.0, 1.25, 1.0, 1.25, 2.0)^T.$$

Nakoniec môžeme vypočítať efektívnosť rovnomerného návrhu vzhľadom na optimálny návrh

$$D_{eff}(\xi) = \left( \frac{\det M(\xi)}{\det M(\xi_D)} \right)^{1/2} = 0.7071068.$$

Pre úplnosť ešte dodávame informačnú maticu optimálneho návrhu

$$M(\xi_D) = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 0.5 \end{pmatrix}.$$

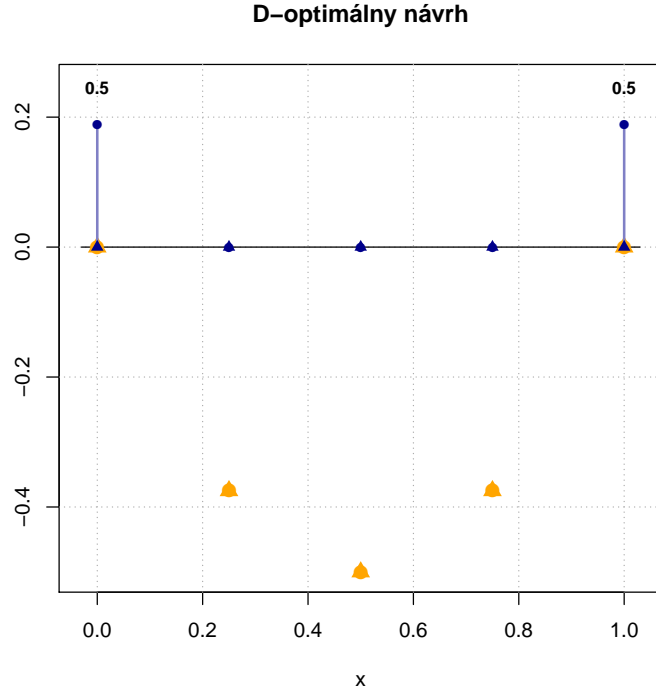
Prišli sme k záveru, že pri lineárnom modeli s dvomi neznámymi parametrami je optimálne podať dávky pacientom tak, že polovici z nich pridáme najvyššiu dávku a druhej polovici zase najnižšiu. Zároveň sme sa presvedčili o tom, že je splnená veta o ekvivalencii, t.j. že maximum variančnej funkcie optimálneho návrhu je rovné počtu neznámych parametrov. Z hľadiska D-efektivity sme prišli k zisteniu, že náš pôvodný návrh pridelit všetkým pacientom rovnaké dávky bol “dobrý” približne na 71% v porovnaní s D-optimálnym návrhom.

V praxi sa však môže vyskytnúť situácia, pri ktorej budeme mať k dispozícii nepárny počet pacientov. V tom prípade by sme nevedeli rozdeliť pacientov na polovicu. Predpokladajme, že  $N = 15$  a rozhodli sme sa pridelit pacientom dávky nasledovne:

$$\xi = \begin{Bmatrix} 0 & 0.25 & 0.5 & 0.75 & 1 \\ 7 & 0 & 0 & 0 & 8 \end{Bmatrix},$$

čo je v tomto prípade najlepší možný exaktný návrh. V porovnaní s D-optimálnym návrhom by efektivita u exaktného návrhu vyšla 0.9977753. Ak by sme teda vo všeobecnosti pri experimente mali nepárny počet pokusov, s exaktným návrhom by sme dosiahli skoro rovnakú účinnosť ako s aproximátnym D-optimálnym návrhom.





**Obr. 1:** Grafické znázornenie D-optimálneho návrhu pre lineárny model s dvomi parametrami (zdroj: *vlastné spracovanie v softvéri R*)

### 2.2.5 Návrh s ohraničením

V klinickej praxi je testovanie nového lieku finančne náročné a preto máme často k dispozícii len obmedzený počet pacientov, obmedzený počet maximálnych zlyhaní, t.j. pokusov, pri ktorých má pacient vedľajšie účinky alebo mu daná dávka nezabrala, atď. Preto budeme v našej práci hľadať aj optimálne návrhy, u ktorých budeme predpokladať isté ohraničenie na zdroje. Informácie sme prebrali z [5].

V tejto kapitole budeme pracovať s takým tvarom exaktného návrhu, ktorý priamo reprezentuje počty meraní v danom bode, teda  $w = (w_1, \dots, w_n)$ , pričom  $w_i$  sú nezáporné celé čísla. Na výpočet optimálneho návrhu pri  $k$  rôznych ohraničenach na zdroje môžeme použiť vzťah

$$\sum_{i=1}^n a_{ji}(x_i)w_i \leq b_j, \quad j = 1, \dots, k \quad (2.15)$$

kde  $a_{ji}(x_i) \geq 0$  označuje spotrebu j-teho zdroja pri pokuse v bode  $x_i$  a  $b_j > 0$  označuje ohraničenie j-teho zdroja. Množina  $\Xi$  označuje všetky realizovateľné návrhy pri

daných ohraničeniach experimentu. Tento typ ohraničení zahŕňa mnohé praktické situácie: uvedieme si niektoré z nich.

Ako sme už spomínali pri testovaní máme často k dispozícii len vyhradený počet pacientov

$$w_1 + \dots + w_n \leq N. \quad (2.16)$$

Pri štandardnom ohraničení (2.16) na celkový počet pacientov  $N$ , pričom predpokladáme, že náklady na každého jedného pacienta sú vo výške 1 peňažnej jednotky dostaneme

$$a_{1i} = 1, \forall i \in \{1, \dots, n\}; \quad b_1 = N.$$

Ak máme ale limitovaný celkový rozpočet, povedzme vo výške  $S$  peňažných jednotiek, a náklady na pacientov, ktorí sa zúčastňujú testovania sú rôzne potom

$$a_{11} = s_1, \dots, a_{1n} = s_n; \quad b_1 = S,$$

kde  $s_1, \dots, s_n$  predstavujú náklady v jednotlivých bodoch  $x_i$  návrhu experimentu.

Podobne by sme dostali prvky  $a_{ji}$  a  $b_j$ , ak by sme uvažovali návrh, pri ktorom by sme mali určený maximálny počet pokusov, ktoré by sme mohli uskutočniť v každom bode návrhu. Alebo môžeme mať návrh s marginálnym ohraničením, t.j. taký u ktorého nesmieme presiahnuť počet pokusov v určitej podmnožine  $\mathcal{X}_j$  pre  $j = 1, \dots, J$ , kde  $\mathcal{X} = \mathcal{X}_1 \cup \dots \cup \mathcal{X}_J$  reprezentuje všetky body návrhu. Z hľadiska testovania nového lieku, by sme mohli množinu  $\mathcal{X}$  rozložiť do disjunktných podmnožín podľa pacientov tak, že ich začleníme do jednotlivých podmnožín na základe veku, dĺžky trvania choroby, pohlavia alebo krajiny z ktorej pochádzajú, pričom nesmieme prekročiť maximálny počet pacientov v danej podmnožine.

V našej bakalárskej práci budeme pri hľadaní optimálneho návrhu s ohraničením používať všeobecný maticový zápis nerovnosti (2.15) v tvare

$$Aw \leq b, \quad A \in \mathbb{R}^{k \times n}, b \in \mathbb{R}^k. \quad (2.17)$$

Konkrétne budeme pri výpočtoch používať balík `OptimalDesign` dostupný v štatistickom softvéri R [13], funkcie `od_RC`, `od_MISOCP` a `od_AQUA`, pri ktorých potrebujeme poznať maticu  $A$  a vektor  $b$ .

Funkcia `od_RC` je heuristický algoritmus [8], ktorý počíta exaktné návrhy za predpokladu ohraničení na zdroje. Funkcia `od_MISOCP` [15] vie narábať s ľubovoľnými line-

árnymi ohraňieniami a vie počítať aproximatívne aj exaktné návrhy. Na použitie tejto funkcie ale potrebujeme mať nainštalovaný externý solver Gurobi [7], ktorý dokáže nájsť najlepšie riešenie zo všetkých prípustných možností. Algoritmus od\_AQUA [6] vie počítať efektívne exaktné návrhy pri lineárnych ohraňieniach, pričom takisto využíva výhody solveru Gurobi.

### 3 Nelineárny model

Na modelovanie vzťahu medzi dávkou lieku a jeho odozvou sa v klinických skúškach častejšie používajú zložitejšie modely a to tzv. nelineárne modely. Postup pri hľadaní optimálnych návrhov a odhadovaní neznámych parametrov u takýchto modelov je oveľa náročnejší. V tejto kapitole si popíšeme základné vlastnosti nelineárneho modelu a následne si uvedieme príklady troch takýchto modelov, ktoré sa prevažne používajú pri testovaní nového liečiva. Nasledujúce poznatky sme čerpali z [14], [3].

Uvažujme nelineárny model v tvare:

$$y(x) = \eta(x, \theta) + \varepsilon, \quad (3.1)$$

pričom  $y(x) \in \mathbb{R}^N$  je pozorovanie v danom  $x$ ,  $\varepsilon \in \mathbb{R}^N$  a  $\theta \in \Theta$  je vektor neznámych parametrov, kde  $\Theta$  (tzv. parametrický priestor) je kompaktná množina v  $\mathbb{R}^m$ .

Funkcia  $\eta(x, \theta)$  vyjadruje očakávanú hodnotu  $Y$  pri danom vektore  $x$ , kde  $Y$  je náhodná premenná s daným rozdelením pravdepodobnosti, t.j.

$$E(Y(x)) = \eta(x, \theta). \quad (3.2)$$

Podobne ako v lineárnom modeli, odhad  $\hat{\theta}$  vieme získať metódou najmenších štvorcov:

$$\hat{\theta} = \arg \min_{\theta \in \Theta} \sum_{k=1}^N [y_k - \eta(x_k, \theta)]^2. \quad (3.3)$$

V tomto prípade funkciu  $\eta(x, \theta)$  nevieme napísať ako súčin funkcií  $f(x)$  a vektora  $\theta$  ako u lineárneho modelu, preto ju aproximujeme pomocou Taylorovho rozvoja v okolí bodu  $\theta_t$ :

$$\eta(x, \theta) \cong \eta(x, \theta_t) + (\theta - \theta_t)^T f(x, \theta_t), \quad (3.4)$$

kde

$$f(x, \theta) = \frac{\partial \eta(x, \theta)}{\partial \theta} \quad (3.5)$$

je gradient funkcie  $\eta(x, \theta)$  v  $\theta$ .

Ak  $\theta$  je “dostatočne blízko” bodu  $\theta_t$ , potom na pravej strane výrazu (3.3) dostaneme:

$$\sum_{k=1}^N \left[ y_k - \eta(x_k, \theta_t) - (\theta - \theta_t)^T f(x_k, \theta_t) \right]^2. \quad (3.6)$$

Ďalšia metóda, ktorou by sme získali odhad parametra  $\theta$  je metóda maximálnej vierohodnosti. Nech hustota pozorovania  $y_k$  (alebo rozdelenie pravdepodobnosti v

diskrétnom prípade) je  $p(y_k|x_k, \theta)$ . Logaritmus funkcie vierohodnosti pre  $N$  nezávislých pozorovaní má tvar

$$L_N(\theta) = \frac{1}{N} \sum_{k=1}^N \ln p(y_k|x_k, \theta) \quad (3.7)$$

Potom hodnotu vektora  $\hat{\theta}$  dostaneme ako riešenie maximalizačnej úlohy:

$$\hat{\theta} = \arg \max_{\theta \in \Theta} L_N(\theta). \quad (3.8)$$

Rovnakým spôsobom by sme taktiež mohli postupovať v prípade lineárneho modelu. Podrobné odvodenie nájde čitateľ v [11].

### Informačná matica a optimálny návrh v nelineárnom modeli

Na rozdiel od lineárneho modelu informačná matica v nelineárnych modeloch závisí aj od hodnôt parametra  $\theta$ , ktoré nepoznáme [4], t.j.

$$M(\xi, \theta) = \sum_{x \in \mathcal{X}} M(x, \theta) \xi(x) = \sum_{x \in \mathcal{X}} f(x, \theta) f^T(x, \theta) \xi(x). \quad (3.9)$$

Vo všeobecnosti nie je možné optimalizovať informačnú maticu nelineárneho modelu nezávislú od  $\theta$ . Existujú rôzne prístupy k riešeniu závislosti od neznámych parametrov. V našej práci predpokladáme, že poznáme hodnotu parametra  $\theta$  (ozn.  $\theta^*$ ). Teda položíme  $\theta = \theta^*$ . Týmto postupom dostaneme tzv. lokálne optimálny návrh.[4]

Pri hľadaní lokálne D-optimálneho návrhu v nelineárnych modeloch s  $m$ -parametrami je počet bodov návrhu, podľa Caratheodoryho vety, niekde medzi  $m$  a  $m(m+1)/2$ . Ak je počet bodov návrhu pre  $m$ -parametrový model presne  $m$ , hovoríme že návrh je minimálne nasýtený. Pre takýto D-optimálny návrh, ktorý je minimálne nasýtený, platí, že všetkým bodom návrhu prislúchajú rovnaké váhy. V dôsledku toho je potrebné určiť len  $m$  bodov návrhu.[12]

## 3.1 Logistický model

Na začiatok si uvedieme základný dvojparametrový logistický model používaný na modelovanie premenných binárnej odozvy. Očakávaná odozva nie je modelovaná priamo, ale je určená pravdepodobnosťou určitého výsledku.

Uvažujme náhodnú premennú  $Y$ , ktorá má alternatívne rozdelenie, t.j. nadobúda hodnoty  $Y = 1$ , pričom táto hodnota predstavuje úspech a  $Y = 0$  predstavuje zlyhanie.

Úspech v tomto prípade znamená, že naše pozorovanie má požadovanú vlastnosť (napr. skúmame efektívnosť nového lieku, vedľajšie účinky, atď.) Vektor  $x \in \mathbb{R}^m$  je vektor nezávislých premenných, ktoré vyjadrujú dávky testovaného lieku. Potom logistický model je definovaný nasledovne [14]:

$$P(Y(x) = 1) = \frac{\exp(f^T(x)\theta)}{1 + \exp(f^T(x)\theta)}, \quad \theta \in \mathbb{R}^m, \quad (3.10)$$

kde  $P(Y(x) = 1)$  predstavuje pravdepodobnosť toho, že naše pozorovanie má uvažovanú vlastnosť pri danom vektore  $x$ , a je splnená podmienka  $0 \leq P(Y(x) = 1) \leq 1$ .

V prípade, že vektor  $f^T(x)$  obsahuje iba jednu premennú a intercept, z (3.10) dostaneme logistický model pre dva parametre

$$P(Y(x) = 1) = \frac{\exp(\theta_0 + \theta_1 x)}{1 + \exp(\theta_0 + \theta_1 x)}. \quad (3.11)$$

Vyššie uvedený tvar logistického modelu sa používa pri teoretických výpočtoch, v bežných medicínskych aplikáciách sa uprednostňuje iný tvar modelu:

Nech  $Y$  je opäť premenná odozvy a vektor  $x$  obsahuje iba jedinú premennú. Logistická regresná funkcia je potom definovaná ako

$$P(Y(x) = 1) = \frac{\exp\left(\frac{x-\mu}{\sigma}\right)}{1 + \exp\left(\frac{x-\mu}{\sigma}\right)}, \quad (3.12)$$

kde  $\mu \in \mathbb{R}$  a  $\sigma > 0$ . Parameter  $\mu$  zodpovedá hodnote  $x$ , pre ktorú je pravdepodobnosť, za podmienky, že odozva nastala, rovná 0.5, t.j.  $P(Y(\mu) = 1) = 0.5$ . V klinickej praxi je dávka testovaného lieku označovaná ako kontrolná premenná, zatiaľčo  $\mu$  je tzv. účinná dávka, označuje sa aj ako  $ED_{50}$ , teda predstavuje minimálnu dávku, ktorá dosahuje aspoň polovicu maximálneho účinku lieku. Druhý parameter nemá praktický význam, súvisí so strmou logistickej funkcie. Čím menšie hodnoty  $\sigma$  nadobúda, tým je krivka strmšia. Ak predpokladáme, že parameter  $\theta_1$  v klasickej parametrizácii logistického modelu nadobúda len kladné hodnoty, potom obidve interpretácie modelu sú ekvivalentné a platí  $\mu = -\frac{\theta_0}{\theta_1}$  a  $\sigma = \frac{1}{\theta_1}$ .

Odhad parametrov u týchto dvoch modelov môžeme získať metódou maximálnej vierohodnosti, ale iba v prípade, že pozorovanie spĺňa určité podmienky. Potrebujeme také pozorovanie, pri ktorom hodnoty  $x$  majú odozvu  $Y = 0$  a  $Y = 1$ , t.j. hodnoty  $x$  sa musia prekrývať. Podrobnejšie informácie nájde čitateľ v [14].

Pri našom experimente budeme pracovať s dvojparametrovým logistickým modelom (3.11), pre ktorý si následne uvedieme aj jeho tvar gradientu  $f(x, \theta)$ , keďže pre

naše výpočty potrebujeme poznať maticu plánu  $F$ :

$$f(x, \theta) = \left( \frac{\exp(\theta_0 + \theta_1 x)}{(1 + \exp(\theta_0 + \theta_1 x))^2}, \frac{x \exp(\theta_0 + \theta_1 x)}{(1 + \exp(\theta_0 + \theta_1 x))^2} \right)^T \quad (3.13)$$

### 3.1.1 Štvorparametrový logistický model

V tejto časti našej bakalárskej práce rozšírime základný dvojparametrový logistický model o ďalšie parametre. Definujeme štvorparametrový logistický model, ktorý má tvar [12]:

$$P(Y(x) = 1) = \theta_0 + \frac{\theta_1 - \theta_0}{1 + \exp(\theta_3(\ln x - \theta_2))}, \quad (3.14)$$

s parametrami  $(\theta_0, \theta_1, \theta_2, \theta_3)$ . Parametre  $\theta_0$  a  $\theta_1$  predstavujú odozvy pri nulovej a “nekonečnej” koncentrácii, v uvedenom poradí,  $\theta_2$  je logaritmus koncentrácie, ktorý udáva odozvu v strede medzi parametrami  $\theta_0$  a  $\theta_1$ , a  $\theta_3$  je parameter sklonu.

Gradient tejto logistickej krivky je daný nasledovne:

$$f(x, \theta) = \begin{pmatrix} 1 - \frac{1}{1 + \exp(\theta_3(\ln x - \theta_2))} \\ \frac{1}{1 + \exp(\theta_3(\ln x - \theta_2))} \\ \frac{(\theta_1 - \theta_0) \exp(\theta_3(\ln x - \theta_2)) \theta_3}{(1 + \exp(\theta_3(\ln x - \theta_2)))^2} \\ \frac{-(\theta_1 - \theta_0) \exp(\theta_3(\ln x - \theta_2)) (\ln x - \theta_2)}{(1 + \exp(\theta_3(\ln x - \theta_2)))^2} \end{pmatrix} \quad (3.15)$$

## 3.2 EMAX model

EMAX model sa podobne ako logistický model taktiež používa na modelovanie binárnej odozvy, je flexibilný a ovplyvňuje vzťah dávka-odozva lieku. Pri podaní väčšej dávky sa zvyšuje aj účinok lieku, pričom maximálny účinok sa dosahuje asymptoticky pri vysokých dávkach. Poznatky sme čerpali z [2] a [14].

Nech  $Y$  je premenná binárnej odozvy a  $x$  vyjadruje dávku lieku. EMAX model definujeme ako

$$P(Y(x) = 1) = \theta_0 + \frac{x\theta_1}{x + \theta_2}, \quad (3.16)$$

kde  $\theta_0$  vyjadruje placebo efekt pri  $x = 0$ ,  $\theta_1$  predstavuje maximálny efekt dávky  $x$  na  $Y$  a  $\theta_2$  určuje dávku, ktorá má polovicu maximálneho efektu. V súlade so značením

(3.2) gradient funkcie  $\eta(x, \theta)$  má tvar

$$f(x, \theta) = \left( 1, \frac{x}{x + \theta_2}, -\frac{\theta_1 x}{(x + \theta_2)^2} \right)^T. \quad (3.17)$$

Následne sa dajú vypočítať D-optimálne návrhy modelu, ktoré čitateľ môže nájsť [2].

### 3.3 Log-lineárny model

Log-lineárny model nám umožňuje testovať aj dávky, ktoré sa blížia k nekonečnu, pretože vytvára v logaritme neohraničený efekt. Tento model opisuje vzťah medzi dávkou lieku a jeho účinkom, kde účinok rastie lineárne v logaritme dávky [2], t.j.

$$P(Y(x) = 1) = \theta_0 + \theta_1 \ln(x + \theta_2), \quad (3.18)$$

kde  $x$  predstavuje dávku lieku,  $\theta_0$  vyjadruje placebo efekt,  $\theta_1$  vyjadruje prírastok daného logaritmu a  $\theta_2$  je aditívna konštanta, ktorá predchádza problému ak uvažujeme placebo efekt. Gradient funkcie  $\eta(x, \theta)$  pre log-lineárny model je tvaru

$$f(x, \theta) = \left( 1, \ln(x + \theta_2), \frac{\theta_1}{(x + \theta_2)} \right)^T. \quad (3.19)$$



## 4 Realizácia experimentu

V predchádzajúcich kapitolách sme sa oboznámili s teóriou, ktorú teraz aplikujeme na konkrétnych príkladoch z praxe. Pomocou nelineárnych modelov sa budeme snažiť získať čo najlepšie optimálne návrhy pre podávanie liečiva, pričom nás bude zaujímať aj efektívnosť našich návrhov, pomocou ktorej vieme usúdiť na koľko presný sme boli pri odhade neznámych parametrov.

### 4.1 A- a D-optimálne návrhy pre dvojparametrový logistický model

Nech  $Y(x)$  je náhodná premenná, ktorá môže nadobúdať iba dve hodnoty, a to v závislosti od toho, či príslušná dávka  $x$  je účinná alebo nie, t.j.

$$Y(x) = \begin{cases} 1 & \text{dávka } x \text{ je účinná} \\ 0 & \text{dávka } x \text{ je neúčinná} \end{cases}$$

Pravdepodobnosť, že dávka  $x$ , ktorú sme podali pacientovi je účinná, bude opisovať logistická krivka tvaru

$$P[Y = 1|X = x] = \frac{\exp(\theta_0 + x\theta_1)}{1 + \exp(\theta_0 + x\theta_1)}. \quad (4.1)$$

Pre jednoduchosť označme výraz (4.1) ako  $p_e(x)$ .

Okrem efektivity nás bude súčasne zaujímať či podaná dávka nie je pre pacienta toxická. Uvažujme preto náhodnú premennú  $Z(x)$ , ktorá nám bude poskytovať informáciu o toxicite dávky  $x$ ,

$$Z(x) = \begin{cases} 1 & \text{dávka } x \text{ má vedľajšie účinky} \\ 0 & \text{dávka } x \text{ nemá vedľajšie účinky} \end{cases}$$

Rovnako ako efektívnosť aj toxicitu dávok budeme modelovať pomocou logistickej krivky

$$P[Z = 1|X = x] = \frac{\exp(\theta_0 + x\theta_1)}{1 + \exp(\theta_0 + x\theta_1)}. \quad (4.2)$$

Pravdepodobnosť, že dávka sa ukáže ako toxická (4.2) budeme označovať  $p_t(x)$ .

Keďže pracujeme s nelineárnym modelom a informačná matica závisí od hodnôt parametra  $\theta$ , zvolíme pre obidva logistické modely štartovacie parametre  $\theta^*$ . Pre efektívnosť  $\theta_e^* = (-10, 0.2)$  a pre logistickú krivku toxicity  $\theta_t^* = (-20, 0.2)$ .

Uvažujme, že nový liek môžeme testovať dávkami v rozpätí od 1 do 150 mg. Pri našom experimente nás budú zaujímať tie dávky, u ktorých sa prejavia vedľajšie účinky a súčasne také, ktoré budú neúčinné, čo môžeme zapísať ako

$$p_f(x) = 1 - (1 - p_t(x))p_e(x).$$

Veličina  $p_f(x)$  teda udáva pravdepodobnosť zlyhania pokusu pre dávku  $x$ .

Ďalej uvažujme, že pre tento experiment máme k dispozícii len  $N = 100$  pacientov, ktorým podávame dávky z množiny  $(x_1, \dots, x_{150})$ , teda pre exaktný návrh  $w = (w_1, \dots, w_{150})$  musí platiť

$$\sum_{i=1}^{150} w_i \leq 100,$$

a počet pokusov, pri ktorých dávka nezabrala alebo mal pacient vedľajšie účinky, môže byť maximálne 10%, t.j.

$$\sum_{i=1}^{150} p_f(x_i)w_i \leq 10.$$

Matica  $A$  a vektor  $b$  pre ohraničenia zo vzorca (2.17) budú mať teda tvar

$$A = \begin{pmatrix} 1 & 1 & \dots & 1 \\ p_f(1) & p_f(2) & \dots & p_f(150) \end{pmatrix}, \quad b = \begin{pmatrix} 100 \\ 10 \end{pmatrix}. \quad (4.3)$$

Našou úlohou je, za daných predpokladov, nájsť exaktný optimálny návrh. V našej práci budeme hľadať také návrhy, ktoré sú optimálne vzhľadom na kritériá A a D. Na to, aby sme ich našli, potrebujeme vyriešiť optimalizačný problém

$$\max_{\xi \in \Xi} \Phi(M(\xi, \theta))$$

pri ohraničení  $Aw \leq b$ .

Pri optimalizovaní úlohy použijeme štatistický softvér R [13]. Spustením funkcie `od_RC` sme dostali D-optimálny návrh:

```
$w.best
 [1] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 [18] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 [35] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 [52] 0 0 0 0 0 0 0 0 0 0 0 0 1 3 4 17 11 27
 [69] 7 15 1 0 0 0 0 0 0 0 0 0 0 0 1 2 4 0
 [86] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 [103] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 [120] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
```

```
[137] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 7
```

\$M.best

```
          G1          G2
G1 4.534934e-03 1.800175e-05
G2 1.800175e-05 4.366743e-04
```

\$Phi.best

```
[1] 0.001407112
```

Výstupom funkcie `od_RC` je 13-bodový exaktný D-optimálny návrh, ktorý spĺňa všetky vyššie uvedené predpoklady a podľa ktorého môžeme nastaviť optimálne dávkovanie lieku pre pacientov. Súčasne sme dostali aj informačnú maticu návrhu  $M(\xi, \theta)$  a hodnotu kritéria D-optimality.

Teraz by sme chceli zistiť na koľko “dobrý” je náš návrh pri predpokladanom 10% ohraničení. Na základe definície pre D-efektivitu (2.13) porovnáme optimálny exaktný návrh s aproximátnym návrhom, u ktorého neuvažujeme žiadne ohraničenie. Aproximátny návrh vieme získať použitím funkcie `od_REX` [9], [10]. Po dosadení optimálnych hodnôt oboch návrhov do (2.13) nám efektivita vyšla 0.3099658. Teda, čo sa týka presnosti odhadov neznámych parametrov, náš návrh by bol oproti optimálnemu “dobrý” len na 30%. Takáto situácia ale v klinických skúškach nastáva často, pretože musíme predovšetkým dbať na bezpečnosť pacientov a ďalšie etické hľadiská.

Pri rovnakých predpokladoch vieme spustením funkcie `od_RC` získať aj A-optimálny návrh pre náš experiment:

\$w.best

```
[1] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
[20] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
[39] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
[58] 0 0 0 1 8 3 10 23 7 10 9 1 1 0 0 0 0 0
[77] 0 0 1 0 0 0 2 2 0 5 8 0 4 0 0 0 0 0
[96] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
[115] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
[134] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 5
```

Vidíme, že pri A-optimálnosti sme dostali úplne iný návrh, pri ktorom by sme pacientom podávali odlišné dávky. Pomocou definície pre A-efektivitu (2.14) nám efektivita u tohto návrhu vyšla 0.1492046, čo je približne o polovicu menej ako v prípade D-optimálneho návrhu. Vidíme, že ohraničenie na počet zlyhaní vo väčšej miere vplýva na efektivitu

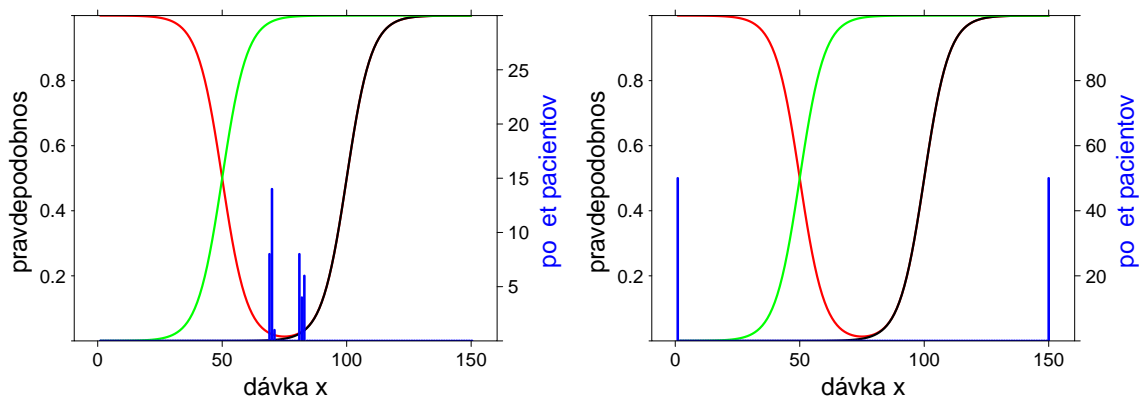
návrhu pre kritérium A-optimality.

Uvažujme extrémne prípady ohraničení na počet zlyhaní, t.j. pravdepodobnosť zlyhania nesmie presiahnuť 1 percento alebo pripustíme až 100%-né zlyhanie. Pomocou funkcie `od_RC` a so vstupným vektorom  $b = (100, 1)^T$  dostaneme exaktný A- a D-optimálny návrh pri 1%-nom ohraničení na počet zlyhaní, a pri vektore  $b = (100, 100)^T$  získame návrhy, pri ktorých dovoľíme úplné zlyhanie. Jednotlivé návrhy sú znázornené na Obr.2 a Obr.3 spolu s funkciami popisujúcimi toxicitu (čierna) a účinnosť (zelená), ktoré vytvárajú akési terapeutické okno pre podávanie dávok, a červená krivka ohraničuje pravdepodobnosť zlyhania. V tabuľke 1 sú uvedené hodnoty pre efektivitu exaktných návrhov. Vidíme, že v prípade A-optimality by náš návrh nedosiahol maxi-

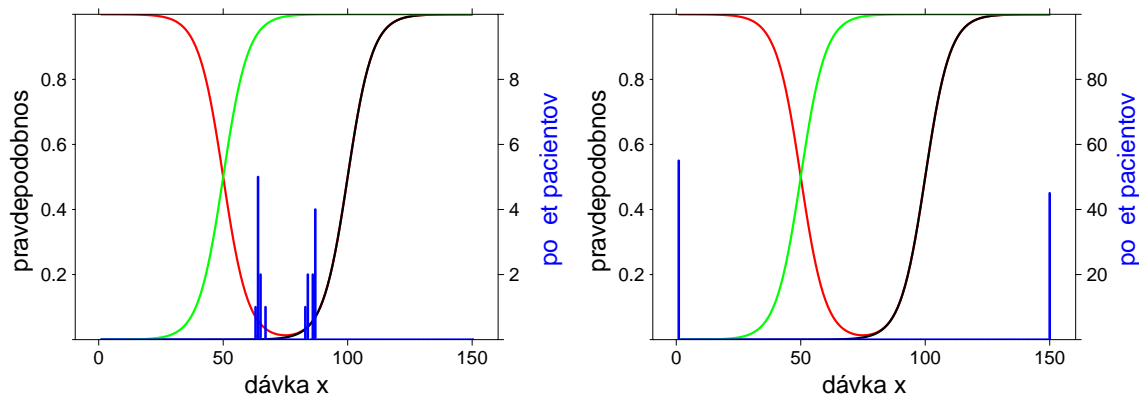
$p_f(x)$	A-efektivita	D-efektivita
1%	0.007326774	0.03356042
100%	0.9999999	1

**Tabuľka 1:** Hodnoty efektivity

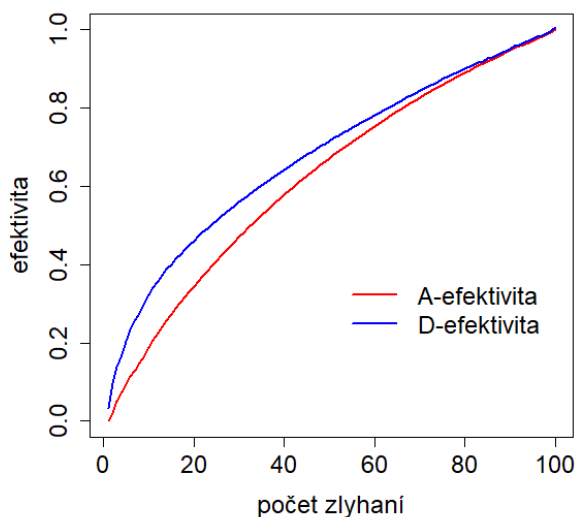
málnu možnú efektivitu ani keby sme pripustili 100%-nú pravdepodobnosť zlyhania. Táto odchýlka je spôsobená tým, že počítame efektivitu vzhľadom na aproximatívny návrh, čiže teoretické optimum, ktoré sa pri danom počte pacientov  $N = 100$  nemusí dať dosiahnuť.



**Obr. 2:** Grafické znázornenie D-optimálneho návrhu (modrá) pri 1%-nej (vľavo) a 100%-nej pravdepodobnosti zlyhania (červená) spolu s účinnosťou (zelená) a toxicitou (čierna)  
(zdroj: vlastné spracovanie v softvéri R)



**Obr. 3:** Grafické znázornenie A-optimálneho návrhu (modrá) pri 1%-nej (vľavo) a 100%-nej pravdepodobnosti zlyhania (červená) spolu s účinnosťou (zelená) a toxicitou (čierna)  
(zdroj: vlastné spracovanie v softvéri R)



**Obr. 4:** Závislosť efektivity exaktných návrhov od počtu maximálnych zlyhaní pri dvojparametrovom logistickom modeli  
(zdroj: vlastné spracovanie v softvéri R)

Ostatné hodnoty pre efektivitu pri rôznych počtoch zlyhaní sú znázornené na Obr.4. Optimálne hodnoty pre D-efektivitu sme získali spustením funkcie `od_MISOCP`. Pri A-optimálnosti funkcia `od_MISOCP` v istých bodoch nenašla riešenie optimalizačného problému v dostatočne krátkom čase. Nie je ale vylúčené, že ak by sme jej dali viac času našla by optimálne návrhy, to ale môže byť kludne aj pol hodina na jeden návrh. Preto

sme sa rozhodli na výpočet A-optimálnych hodnôt použiť algoritmus od\_AQUA. Obidve funkcie mali na výpočet stanovený maximálny čas 15 sekúnd. Pri A- aj D- optimálnych návrhoch si môžeme všimnúť, že čím menej zlyhaní od experimentu požadujeme, tým menej efektívny návrh sme schopní dosiahnuť, pričom v každom bode zlyhania dosiahne D-optimálny návrh vyššiu efektivitu.

## 4.2 A- a D-optimálne návrhy pre štvorparametrový logistický model

Účinnosť a toxicitu dávok budeme teraz modelovať pomocou štvorparametrového logistického modelu

$$p_e(x) = p_t(x) = \theta_0 + \frac{\theta_1 - \theta_0}{1 + \exp(\theta_3(\ln x - \theta_2))},$$

a ako štartovacie hodnoty pre neznáme parametre  $\theta$  zvolíme  $\theta_e^* = (0, 0.6, 3.5, -4)$  a  $\theta_t^* = (0, 0.4, 4.5, -5)$ .

Rovnako ako pri predchádzajúcom modeli predpokladáme, že máme k dispozícii 100 pacientov a pravdepodobnosť zlyhania môže byť maximálne  $u$ , teda  $p_f(x) \leq u$ . Porovnáme A- a D-optimálne návrhy za predpokladu, že pri experimente pripustíme 50%-né zlyhanie, teda pre  $u = 50$ . Pri výpočte exaktných návrhov sme spustili funkcie od\_MISOCP a od\_AQUA s časovým limitom 10 minút. Pre D-optimálnu nám pomocou od\_MISOCP vyšiel návrh:

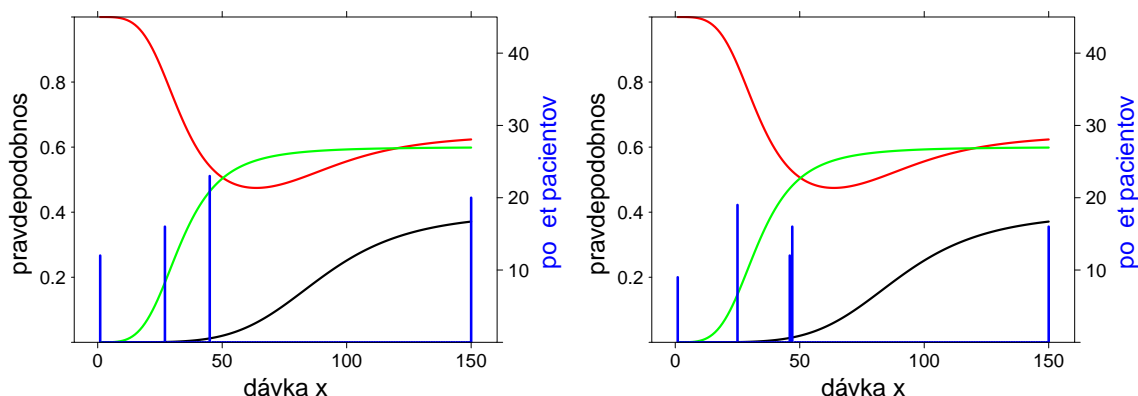
```
$w.best
[1] 12 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
[23] 0 0 0 0 16 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
[45] 23 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
[67] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
[89] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
[111] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
[133] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 20
```

ktorý by z celkových 150 dávok využil pri testovaní len štyri, zatiaľ čo funkcia od\_AQUA sa dopracovala k takému A-optimálnemu návrhu, pri ktorom sa objavila aj piata dávka:

```
$w.best
[1] 9 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
[23] 0 0 19 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
```

[45]	0	12	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
[67]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
[89]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
[111]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
[133]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	16			

Môžeme si všimnúť, že pri oboch návrhoch by sa dávky delili medzi približne rovnaký počet pacientov. Na Obr.5 sú spolu s návrhmi znázornené aj funkcie pre toxicitu, účinnosť a zlyhanie, ktoré ohraničujú terapeutické okno dávkovania.



**Obr. 5:** Grafické znázornenie D- (vľavo) a A-optimálneho návrhu (modrá) pri 50%-nej pravdepodobnosti zlyhania (červená) spolu s účinnosťou (zelená) a toxicitou (čierna) (zdroj: *vlastné spracovanie v softvéri R*)

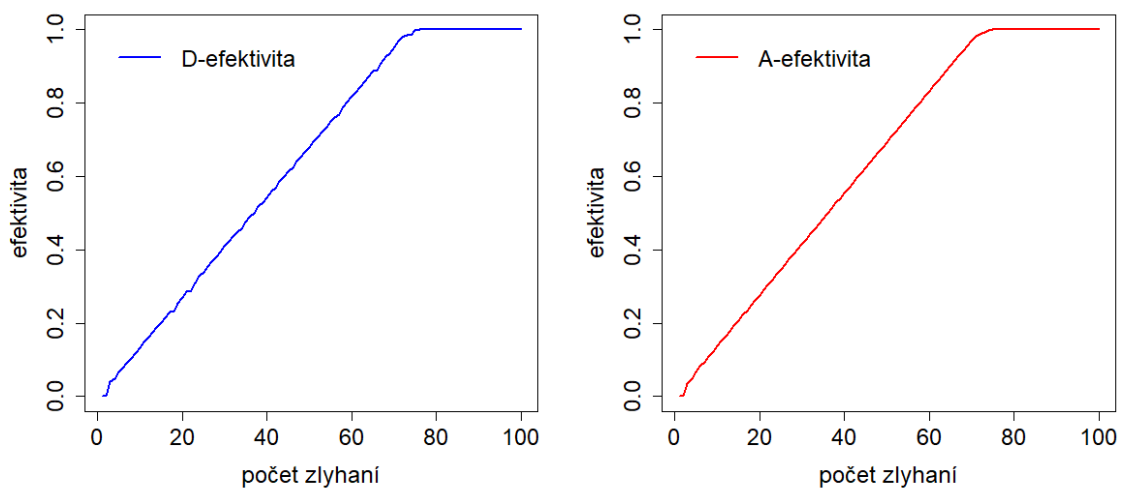
Aby sme mohli usúdiť na koľko presný sme boli pri odhade neznámych parametrov, potrebujeme hodnoty pre A- a D-efektivitu. Pomocou optimálnych hodnôt, ktoré sme dostali výstupom funkcií `od_MISOCP`, `od_AQUA` a zodpovedajúcich aproximatívnych návrhov, nám vyšli hodnoty:

$$\text{D-efektivita} = 0.6827285, \quad \text{A-efektivita} = 0.6935278$$

Náš exaktný D-optimálny návrh by bol teda v porovnaní s príslušným teoretickým optimom “dobrý” na 68%, zatiaľ čo A-optimálny návrh by dosiahol len o 1% väčšiu účinnosť.

Doteraz sme uvažovali, že pravdepodobnosť zlyhania pokusu pre dávku  $x$  nesmie byť väčšia ako 50%. Následne nasimulujeme exaktné návrhy pre všetky možné počty zlyhaní, teda pre  $u = 1, 2, \dots, 100$ . Podobne ako pri dvojparametrovom logistickom modeli, na výpočet D-optimálnych návrhov využijeme funkciu `od_MISOCP` a

na A-optimálne návrhy aplikujeme funkciu `od_AQUA`, keďže aj pri tomto modeli bola A-optimálnosť oveľa náročnejšia a algoritmus `od_MISOCP` sa v stanovenom čase nedopracoval k riešeniu. Funkcie mali na výpočet stanovený limit 15 sekúnd. Výsledné hodnoty efektivity návrhov pri jednotlivých ohraničeniach sú znázornené na Obr.6. Môžeme si všimnúť, že v oboch prípadoch rastie krivka efektivity takmer lineárne, teda efektivita exaktného návrhu je tým väčšia, čím väčší počet zlyhaní pri experimente pripustíme, pričom už pri menej ako 80%-nom ohraničení dosahuje efektivita maximálnu hodnotu.



**Obr. 6:** Závislosť efektivity exaktných návrhov od počtu maximálnych zlyhaní pri štvorparametrovom logistickom modeli  
(zdroj: vlastné spracovanie v softvéri R)

### 4.3 A- a D-optimálne návrhy pre log-lineárny model a EMAX model

Opäť predpokladáme, že  $p_t(x)$  predstavuje pravdepodobnosť toxicity dávok a  $p_e(x)$  pravdepodobnosť, že dávka  $x$  je účinná. Vypočítame exaktné optimálne návrhy vzhľadom na kritériá A a D, ak tieto pravdepodobnosti budú popisovať log-lineárny model a EMAX model. V tejto podkapitole skúmame efektívnosť návrhov iba pre 10%-né ohraničenie na počet zlyhaní s limitovaným počtom pacientov 100. Pre návrh s týmto lineárnym ohraničením má matica  $A$  a vektor  $b$  rovnaký tvar ako v (4.3).



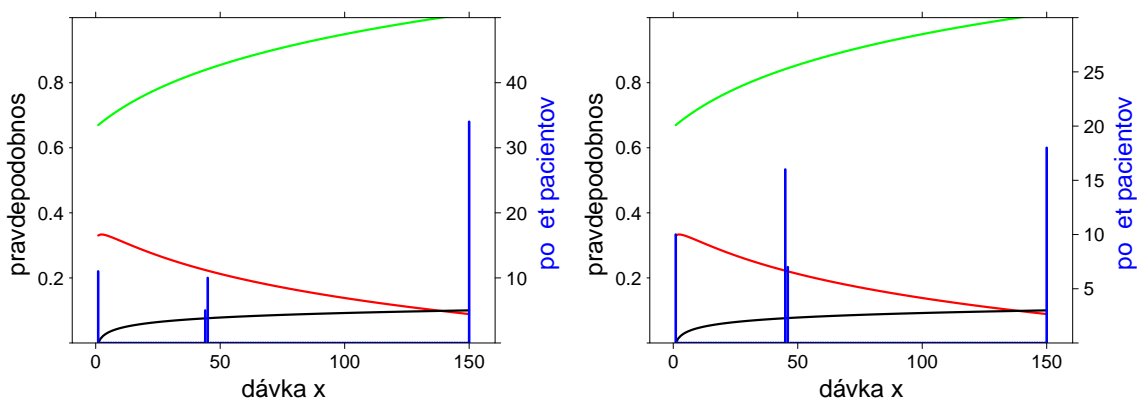
Uvažujme najskôr log-lineárny model

$$p_e(x) = p_t(x) = \theta_0 + \theta_1 \ln(x + \theta_2),$$

pre ktorý sme zvolili začiatočné hodnoty  $\theta_e^* = (0, 0.195, 30)$  a  $\theta_t^* = (0, 0.02, 0)$  s ohraničením pre pravdepodobnosť zlyhania  $p_f(x) \leq 10$ . Pri takýchto predpokladoch pomocou funkcie `od_AQUA` dostaneme D-optimálny návrh:

```
$w.best
 [1] 11  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
 [23]  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  5
 [45] 10  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
 [67]  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
 [89]  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
[111]  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
[133]  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  34
```

ktorý pri porovnaní s aproximatívnym návrhom bez ohraničenia dosiahol efektívitu 0.5258135. Pri tomto modeli má funkcia pravdepodobnosti zlyhania minimum v najvyššej dávke a preto sa algoritmus `od_AQUA` snažil sústrediť čo najviac meraní (pacientov) do tohto bodu. Pre kritérium A-optimality návrh vyšiel taktiež 4-bodový (viď Obr.7), ale nie je až tak výrazný rozdiel medzi počtom meraní v jednotlivých dávkach, hoci najvyššej dávke prislúcha stále najviac pacientov. Z hľadiska odhadnutých parametrov je tento návrh oproti teoretickému optimu “dobrý” približne na 46%.



**Obr. 7:** Grafické znázornenie D- (vľavo) a A-optimálneho návrhu (modrá) pre log-lineárny model s krivkami pre toxicitu (čierna), účinnosť (zelená) a zlyhanie (červená) (zdroj: *vlastné spracovanie v softvéri R*)

Ak by pravdepodobnosti  $p_t(x)$  a  $p_e(x)$  opisoval EMAX model

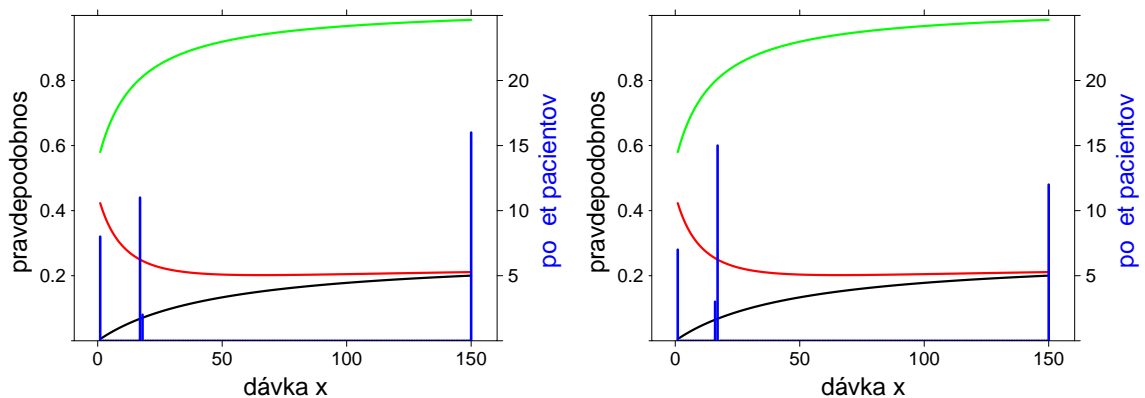
$$p_e(x) = p_t(x) = \theta_0 + \frac{x\theta_1}{x + \theta_2},$$

s odhadnutými parametrami  $\theta_e^* = (0.55, 0.48, 15)$  a  $\theta_t^* = (0, 0.267, 50)$  a s rovnakým 10%-ným ohraničením, spustením funkcie `od_AQUA` dostaneme exaktný D-optimálny návrh, pri ktorom by sme z možných 150 dávok vybrali tieto štyri:

\$w.best

[1]	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11	2	0	0	0	0
[23]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
[45]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
[67]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
[89]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
[111]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
[133]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Rovnako ako v prípade predchádzajúcich nelineárnych modelov, hodnotu D-optimality aproximatívneho návrhu dostaneme použitím funkcie `od_REX`. Na základe optimálnych hodnôt pre exaktný a aproximatívny návrh nám vyšla D-efektivita pri EMAX modeli rovná 0.3524795, z čoho môžeme usúdiť, že pri našich odhadnutých parametroch je D-optimálny návrh vzhľadom na teoretické optimum “dobrý” na 35%. Čo sa týka A-optimality tam vyšiel návrh veľmi podobný (pozri Obr.8), dokonca s takmer identickou efektívnosťou: 0.3571871.



**Obr. 8:** Grafické znázornenie D- (vľavo) a A-optimálneho návrhu (modrá) pre EMAX model s krivkami pre toxicitu (čierna), účinnosť (zelená) a zlyhanie (červená) (zdroj: *vlastné spracovanie v softvéri R*)

Na Obr.7 a Obr.8 si môžeme všimnúť aj to, aké terapeutické okná vytvorili krivky, ktoré ohraničujú toxicitu a účinnosť, porovnať návrhy oboch kritérií a sledovať priebeh funkcie pravdepodobnosti podania dávky, ktorá je neúčinná alebo má vedľajšie účinky.

Vo všeobecnosti by sme mohli povedať, že pre náš experiment sme našli optimálne riešenia pre všetky nelineárne modely, ktoré popisovali pravdepodobnosť podania neúčinnnej dávky alebo dávky s vedľajšími účinkami. Nominálne hodnoty parametra boli zvolené tak, aby zodpovedali situáciám, v ktorých je vhodné veľké terapeutické okno a pre ktoré by dané nelineárne modely reprezentovali pravdepodobnosti účinnosti a toxicity. Pri hľadaní optimálnych návrhov sme pri výpočtoch používali algoritmy `od_RC`, `od_MISOCP` a `od_AQUA`. Prvé dva algoritmy sme používali zväčša v prípade D-optimality a k návrhom sme sa dopracovali takmer okamžite. Problematickejšie boli A-optimálne návrhy, pri ktorých sme museli vyskúšať viacero možností, no nakoniec sa ako najlepšie riešenie ukázalo použiť funkciu `od_AQUA`. Nevylúčili sme však ani fakt, že by sa k návrhom dopracovali aj ostatné algoritmy, len by zrejme potrebovali oveľa dlhší časový limit na výpočet.

## Záver

V tejto bakalárskej práci sme sa venovali navrhovaniu experimentov z klinickej oblasti, pričom sme brali do úvahy lineárne ohraničenia. Skúmali sme reláciu medzi dávkou a odozvou pacienta, ktorú odhadovali nelineárne modely a hľadali sme optimálne návrhy tak, aby pravdepodobnosť podania neúčinnnej alebo toxickej dávky nepresiahla nejakú hodnotu. Efektívnosť získaných návrhov sme počítali vzhľadom na teoretické optimum, čiže aproximatívne návrhy bez ohraničenia.

Práca pozostáva zo 4 kapitol. V prvej kapitole sme zhrnuli postup zavádzania medikamentov do klinickej praxe, ktorý pozostáva zo štyroch fáz podliehajúcich tzv. „Správnej klinickej praxi“. Druhá kapitola definuje lineárny model, pojmy ako aproximatívny a exaktný návrh, informačná matica, kritériá optimality, efektívnosť návrhu a popisuje návrh s lineárnym ohraničením. Tretia kapitola sa zaoberá vlastnosťami nelineárnych modelov, z ktorých sme priblížili EMAX model, log-lineárny model a logistický model s dvomi a štyrmi neznámymi parametrami. Tieto kapitoly vo všeobecnosti predstavujú teoretický základ k realizácii experimentu, ktorý je obsahom štvrtej kapitoly.

Nelineárne modely popisovali pravdepodobnosť toxicity a účinnosti dávok. Neznáme parametre týchto modelov sme potrebovali odhadnúť tak, aby sme vytvorili akési terapeutické okno, v ktorom je možné podávať dávky. Ohraničenie v našom experimente spočívalo v tom, že sme mali limitovaný počet pacientov a pravdepodobnosť nášho zlyhania, t.j. podania neúčinnnej alebo toxickej dávky, nesmela presiahnuť istú hodnotu.

V prvej časti poslednej kapitoly sme pracovali s dvojparametrom logistickým modelom. Hľadali sme exaktné optimálne návrhy vzhľadom na kritériá A a D najskôr pri 10%-nom ohraničení a následne sme zisťovali závislosť medzi počtom zlyhaní a efektívnosťou návrhu. Z výsledkov sme zistili, že efektívnosť rastie so zvyšujúcim sa počtom zlyhaní, čo bolo očakávateľné. Podobným spôsobom sme postupovali v prípade logistického modelu so štyrmi parametrami. Na začiatku sme hľadali optimálne návrhy pre jedno ohraničenie a v ďalšej časti sme počítali efektívnosť návrhov pre rôzne ohraničenia. V prípade A-optimality sme museli vyskúšať viacero algoritmov, aby sme našli čo najlepšie riešenie nášho problému. Nakoniec sa ukázalo, že pri tomto modeli efektívnosť

návrhov rastie lineárne, ale maximum dosiahne už pri menej ako 80%-nom ohraničení.

Na záver sme hľadali D-optimálne návrhy pre log-lineárny a EMAX model, pri ktorých sme uvažovali maximálne 10 zlyhaní a skúmali sme ich efektivitu, ktorá posudzovala presnosť odhadnutých parametrov. Nakoniec sme pridali aj optimálne návrhy vzhľadom na kritérium A.

Samozrejme, práca by sa dala rozšíriť o ďalšie nelineárne modely, napr. trojparametrový logistický model z článku [12] alebo exponenciálny model z [2], ktoré by popisovali pravdepodobnosť zlyhania nášho experimentu. Prípadne by bolo možné, pozrieť sa ako by vyzerali optimálne návrhy a ich efektivita, ak by sme tieto návrhy počítali vzhľadom na iné kritériá optimality.

## Zoznam použitej literatúry

- [1] Atkinson, A.C., Donev, A.D., Tobias, R.D. 2007. *Optimum Experimental Designs, with SAS*. Oxford University Press Inc., New York.
- [2] Dette, H., Kiss, Ch., Bevanda, M., Bretz, F. 2010. *Optimal designs for the EMAX, log-linear and exponential model*. Biometrika, volume 97. 513-518 str.
- [3] Fedorov, V.V., Leonov, S.L. 2014. *Optimal Design for Nonlinear Response Models*. CRC Press, Taylor & Francis Group, Boca Raton.
- [4] Filová, L. 2017. *Štatistické metódy v klinických skúškach*, poznámky k prednáškam. KAMŠ, FMFI UK, Bratislava.
- [5] Filová, L. *Optimal Designs in the Linear Model*. [cit. 11.4.2020] Dostupné na adrese: <http://www.iam.fmph.uniba.sk/ospm/Filova/rm/OptDesign2.html>
- [6] Filová, L., Harman, R. 2020. *Ascent with quadratic assistance for the construction of exact experimental designs*. Computational Statistics, 1-27.
- [7] Gurobi Optimization. 2020. LLC: Gurobi Optimizer Reference Manual. <http://www.gurobi.com>
- [8] Harman, R., Bachratá, A., Filová, L. 2016. *Heuristic construction of exact experimental designs under multiple resource constraints, Applied Stochastic Models in Business and Industry*. Volume 32, pp.3-17
- [9] Harman, R., Filová, L., Richtárik P. 2019. *A randomized exchange algorithm for computing optimal approximate designs of experiments*. Journal of the American Statistical Association, 1-30.
- [10] Harman, R., Jurik, T. 2008. *Computing c-optimal experimental designs using the simplex method of linear programming*. Computational Statistics and Data Analysis 53, 247-254
- [11] Janková, K., Pázman, A. 2011. *Pravdepodobnosť a štatistika*. Univerzita Komenského Bratislava. 128 str.

- [12] Li, G., Majumdar, D. 2007. *D-optimal designs for logistic models with three and four parameters*. Journal of Statistical Planning and Inference, volume 138, Elsevier B.V.
- [13] R Core Team. 2019. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Dostupné na adrese: <https://www.R-project.org/>.
- [14] Roth, K. 2009. *Optimal Design for Dose Finding Studies on Safety and Efficacy*. Dissertation, die Fakultät für Mathematik, der Otto-von-Guericke-Universität, Magdeburg.
- [15] Sagnol, G., Harman, R. 2015. *Computing exact D-optimal designs by mixed integer second order cone programming*. *The Annals of Statistics*. Volume 43, Number 5, pp. 2198-2224.

# Prílohy

## Príloha A.

Výpočty v softvéri R pre A- a D-optimálne návrhy pri dvojparametrovom logistickom modeli.

```
# matica planu
F.logis<-Fx_glm(~x1,c(-10,0.2),glm.model="bin-logit",n.levels=150)

# aproximativny D-opt.navrh a hodnota D-kriteria
w1<-od_REX(F.logis)$w.best
opt1 <- optcrit(F.logis, 100*w1)

#davky
X<-1:150
#theta pre efektivitu
t1<-c(-10,0.2)
#pravdepodobnost, ze liek je ucinnny
pe<-rep(0,150)
for ( i in 1:150) {
  pe[i]<-(exp(t1[1]+t1[2]*X[i]))/(1+exp(t1[1]+t1[2]*X[i])) }

#theta pre toxicitu
t2<-c(-20,0.2)
#pravdepodobnost, ze liek ma vedlajsie ucinky
pt<-rep(0,150)
for ( i in 1:150) {
  pt[i]<-(exp(t2[1]+t2[2]*X[i]))/(1+exp(t2[1]+t2[2]*X[i])) }

# pravdepodobnost, ze liek je neucinnny a ma vedlajsie ucinky
pf<-rep(0,150)
for ( i in 1:150) { pf[i]<-1-(1-pt[i])*pe[i] }

# 100 pacientov, pocet max. zlyhani=10
b10<-c(100,10)
A<-matrix(1,nrow=2,ncol=150)
for (i in 1:150) { A[2,i]<-pf[i] }

Dopt<-od_RC(F.logis,b10,A,t.max=30)
wD<-Dopt$w.best #exaktny D-opt.navrh pri 10 zlyhaniach
od_plot(F.logis,wD,1:150,main.lab="Exaktný D-optimálny návrh",
  y.lab="počet pacientov")
eff_10<-Dopt$Phi.best/opt1

b1<-c(100,1)
Dopt1<-od_RC(F.logis,b1,A,t.max=30)
wD1<-Dopt1$w.best # D-opt.navrh pri 1 zlyhani
eff_1<-Dopt1$Phi.best/opt1
```



```

b100<-c(100,100)
Dopt100<-od_RC(F.logis,b100,A,t.max=30)
wD100<-Dopt100$w.best      # D-opt.navrh pri 100 zlyhaniach
eff_100<-Dopt100$Phi.best/opt1

#library(lattice)
#library(latticeExtra)
#graficke znazornenie
x <- 1:150
data <- data.frame(x,pf,pt,pe,Dopt100$w.best)
obj1 <- xyplot(pf ~ x, data, type = "l" , lwd=3, col="red",ylim=c(0,1),
              scale=list(y=list(tck=0.5)),xlab=list("dávka x",fontsize=25),
              ylab=list("pravdepodobnosť",fontsize=25),
              ylab.right=list("počet pacientov",fontsize=25,col="blue"),
              par.settings=list(layout.widths=list(ylab.axis.padding=3)))
obj2 <- xyplot(pt ~ x, data, type = "l", lwd=3, col="black")
obj3 <- xyplot(pe ~ x, data, type = "l" , lwd=3, col="green")
obj4 <- xyplot(Dopt100$w.best ~ x, data, type = "h", lwd=3.5,
              col="blue",ylim=c(0,30),ylab="")
trellis.par.set(axis.text=list(cex=1.6),ylab=list(col="blue"))
doubleYScale(obj1+obj2+obj3, obj4, add.ylab2 = TRUE,use.style=FALSE)

data <- data.frame(x,pf,pt,pe,Dopt100$w.best)
obj1 <- xyplot(pf ~ x, data, type = "l" , lwd=3, col="red",ylim=c(0,1),
              scale=list(y=list(tck=0.5)),xlab=list("dávka x",fontsize=25),
              ylab=list("pravdepodobnosť",fontsize=25),
              ylab.right=list("počet pacientov",fontsize=25,col="blue"),
              par.settings=list(layout.widths=list(ylab.axis.padding=3)))
obj2 <- xyplot(pt ~ x, data, type = "l", lwd=3, col="black")
obj3 <- xyplot(pe ~ x, data, type = "l" , lwd=3, col="green")
obj4 <- xyplot(Dopt100$w.best ~ x, data, type = "h", lwd=3.5,
              col="blue",ylim=c(0,100),ylab="")
trellis.par.set(axis.text=list(cex=1.6),ylab=list(col="blue"))
doubleYScale(obj1+obj2+obj3, obj4, add.ylab2 = TRUE,use.style=FALSE)

#vektor zlyhani
u<-1:100

# D-optimalne hodnoty pre vsetky zlyhania
Dopt.phi.best<-rep(0,100)
for ( i in 1:100) { b<-c(100,u[i])
  Dopt.phi.best[i]<-od_MISOCP(F.logis,b,A,t.max=15)$Phi.best }
Dopt.phi.best

# vektor pre D-efektivitu
eff_D<-rep(0,100)
for ( i in 1:100) { eff_D[i]<-Dopt.phi.best[i]/opt1 }
eff_D

# aproximativny A-opt.navrh a hodnota A-optimality
w2<-od_REX(F.logis,crit="A")$w.best
opt2<-optcrit(F.logis,100*w2,crit="A")

```

```

Aopt<-od_RC(F.logis,b10,A,t.max=30,crit="A")
wA<-Aopt$w.best # exaktny A-opt.navrh pri 10 zlyhaniach

effA_10<-Aopt$Phi.best/opt2

Aopt1<-od_RC(F.logis,b1,A,t.max=30,crit="A")
wA1<-Aopt1$w.best # exaktny A-opt.navrh pri 1 zlyhani
effA_1<-Aopt1$Phi.best/opt2

Aopt100<-od_RC(F.logis,b100,A,t.max=30,crit="A")
wA100<-Aopt100$w.best # exaktny A-opt.navrh pri 100 zlyhaniach
effA_100<-Aopt100$Phi.best/opt2

# graficke znazornenie
data <- data.frame(x,pf,pt,pe,Aopt1$w.best)
obj1 <- xyplot(pf ~ x, data, type = "l" , lwd=3, col="red",ylim=c(0,1),
  scale=list(y=list(tck=0.5)),xlab=list("dávka x",fontsize=25),
  ylab=list("pravdepodobnosť",fontsize=25),
  ylab.right=list("počet pacientov",fontsize=25,col="blue"),
  par.settings=list(layout.widths=list(ylab.axis.padding=3)))
obj2 <- xyplot(pt ~ x, data, type = "l", lwd=3, col="black")
obj3 <- xyplot(pe ~ x, data, type = "l" , lwd=3, col="green")
obj4 <- xyplot(Aopt1$w.best ~ x, data, type = "h", lwd=3.5,
  col="blue",ylim=c(0,30),ylab="")
trellis.par.set(axis.text=list(cex=1.6),ylab=list(col="blue"))
doubleYScale(obj1+obj2+obj3, obj4, add.ylab2 = TRUE,use.style=FALSE)

data <- data.frame(x,pf,pt,pe,Aopt100$w.best)
obj1 <- xyplot(pf ~ x, data, type = "l" , lwd=3, col="red",ylim=c(0,1),
  scale=list(y=list(tck=0.5)),xlab=list("dávka x",fontsize=25),
  ylab=list("pravdepodobnosť",fontsize=25),
  ylab.right=list("počet pacientov",fontsize=25,col="blue"),
  par.settings=list(layout.widths=list(ylab.axis.padding=3)))
obj2 <- xyplot(pt ~ x, data, type = "l", lwd=3, col="black")
obj3 <- xyplot(pe ~ x, data, type = "l" , lwd=3, col="green")
obj4 <- xyplot(Aopt100$w.best ~ x, data, type = "h", lwd=3.5,
  col="blue",ylim=c(0,100),ylab="")
trellis.par.set(axis.text=list(cex=1.6),ylab=list(col="blue"))
doubleYScale(obj1+obj2+obj3, obj4, add.ylab2 = TRUE,use.style=FALSE)

# A-optimalne hodnoty pre vsetky zlyhania
Aopt.phi.best<-rep(0,100)
for ( i in 1:100) { b<-c(100,u[i])
  Aopt.phi.best[i]<-od_AQUA(F.logis,b,A,t.max=15,crit="A")$Phi.best }
Aopt.phi.best

#vektor A-efektivty
effA<-rep(0,100)
for ( i in 1:100) { effA[i]<-Aopt.phi.best[i]/opt2 }

```

```
#graficke znazornenie A- a D-efektivita
plot(u,effA,type="l",col = "red", lwd = 2,xlab="počet zlyhaní",
      ylab="efektivita", cex.lab=1.5,cex.axis=1.5)
lines(eff_D,type="l",col = "blue", lwd = 2)
legend(x=50,y=0.4,legend=c("A-efektivita","D-efektivita"),
       col=c("red","blue"),box.col="white",lty=1,lwd=2,cex=1.5)
```

## Príloha B.

Výpočty v softvéri R pre A- a D-optimálne návrhy pri štvorparametrovom logistickom modeli.

```
# Matica planu
Fx<-function(theta,X) {
Fx<-matrix(0,nrow=length(X),ncol=4 )
for ( i in 1:length(X) ) {
Fx[i,]<-c(1-1/(1+exp(theta[4]*(log(X[i])-theta[3]))),
1/(1+exp(theta[4]*(log(X[i])-theta[3]))),
(theta[2]-theta[1])*theta[4]*exp(theta[4]*(log(X[i])-theta[3]))/
(1+exp(theta[4]*(log(X[i])-theta[3])))^2),
-(theta[2]-theta[1])*exp(theta[4]*(log(X[i])-theta[3]))*
(log(X[i])-theta[3])/(1+exp(theta[4]*(log(X[i])-theta[3])))^2))
}
return(Fx)
}

F<-Fx(c(0,0.6,3.5,-4),c(1:150))

X<-1:150 # davky
t1<-c(0,0.6,3.5,-4) # theta pre efektivitu
pe<-rep(0,150) # pravdepodobnost, ze liek je ucinnny

for ( i in 1:150) {
pe[i]<-t1[1]+(t1[2]-t1[1])/(1+exp(t1[4]*(log(X[i])-t1[3]))) }

t2<-c(0,0.4,4.5,-5) # theta pre toxicitu
pt<-rep(0,150) # pravdepodobnost, ze liek ma vedlajsie ucinky

for ( i in 1:150) {
pt[i]<-t2[1]+(t2[2]-t2[1])/(1+exp(t2[4]*(log(X[i])-t2[3]))) }

# pravdepodobnost, ze liek je neucinnny a ma vedlajsie ucinky
pf<-rep(0,150)
for ( i in 1:150) { pf[i]<-1-(1-pt[i])*pe[i] }

b<-c(100,50) # 100 pacientov, pocet max. zlyhani=50
A<-matrix(1,nrow=2,ncol=150)
for (i in 1:150) { A[2,i]<-pf[i] }

Dopt<-od_MISOCP(F,b,A,t.max=600)
wD<-Dopt$w.best # D-opt.navrh pri 50 zlyhaniach

Aopt<-od_AQUA(F,b,A,t.max=600,crit="A")
wA<-Aopt$w.best # A-opt.navrh pri 50 zlyhaniach

w1<-od_REX(F)$w.best # D-aproximativny navrh
opt1 <- optcrit(F,w1*100) # D-optimalna hodnota aprox.navrhu

eff_d<-Dopt$Phi.best/opt1 # D-efektivita pri 50 zlyhaniach
```

```

w2<-od_REX(F,crit="A")$w.best          # A-aproximativny navrh
opt2 <- optcrit(F,w2*100,crit="A")     # A-optimalna hodnota aprox.navrhu

eff_a<-Aopt$Phi.best/opt2             # A-efektivita pri 50 zlyhaniach

#vektor zlyhani
u<-1:100

# D-optimalne hodnoty pre vsetky zlyhania
Dopt.phi.best<-rep(0,100)
for ( i in 1:100) { b1<-c(100,u[i])
  Dopt.phi.best[i]<-od_MISOCP(F,b1,A,t.max=15)$Phi.best }

effD<-rep(0,100)
for ( i in 1:100) { effD[i]<-Dopt.phi.best[i]/opt1 }
plot(u,effD,type="l",col = "blue", lwd = 2,xlab="počet zlyhaní",
      ylim=c(0,1),ylab="efektivita",cex.lab=1.5,cex.axis=1.5)
legend(x=0,y=1,legend=c("D-efektivita"),col=c("blue"),
       box.col="white",lwd=2,cex=1.5)

# A-optimalne hodnoty pre vsetky zlyhania
Aopt.phi.best<-rep(0,100)
for ( i in 1:100) { b1<-c(100,u[i])
  Aopt.phi.best[i]<-od_AQUA(F,b1,A,t.max=15,crit="A")$Phi.best }

effA<-rep(0,100)
for ( i in 1:100) { effA[i]<-Aopt.phi.best[i]/opt2 }
plot(u,effA,type="l",col = "red", lwd = 2,xlab="počet zlyhaní",
      ylab="efektivita",ylim=c(0,1),cex.lab=1.5,cex.axis=1.5)
legend(x=0,y=1,legend=c("A-efektivita"),col=c("red"),
       box.col="white",lwd=2,cex=1.5)

#library(lattice)
#library(latticeExtra)
#graficke znazornenie

x <- 1:150
data <- data.frame(x,pf,pt,pe,Dopt$w.best)
obj1 <- xyplot(pf ~ x, data, type = "l" , lwd=3, col="red",ylim=c(0,1),
               scale=list(y=list(tck=0.5)),xlab=list("dávka x",fontsize=25),
               ylab=list("pravdepodobnosť",fontsize=25),
               ylab.right=list("počet pacientov",fontsize=25,col="blue"),
               par.settings=list(layout.widths=list(ylab.axis.padding=3)))
obj2 <- xyplot(pt ~ x, data, type = "l", lwd=3, col="black")
obj3 <- xyplot(pe ~ x, data, type = "l" , lwd=3, col="green")
obj4 <- xyplot(Dopt$w.best ~ x, data, type = "h", lwd=3.5,
               col="blue",ylim=c(0,30),ylab="")
trellis.par.set(axis.text=list(cex=1.6),ylab=list(col="blue"))
doubleYScale(obj1+obj2+obj3, obj4, add.ylab2 = TRUE,use.style=FALSE)

```

```
data <- data.frame(x,pf,pt,pe,Aopt$w.best)
obj1 <- xyplot(pf ~ x, data, type = "l" , lwd=3, col="red",ylim=c(0,1),
              scale=list(y=list(tck=0.5)),xlab=list("dávka x",fontsize=25),
              ylab=list("pravdepodobnosť",fontsize=25),
              ylab.right=list("počet pacientov",fontsize=25,col="blue"),
              par.settings=list(layout.widths=list(ylab.axis.padding=3)))
obj2 <- xyplot(pt ~ x, data, type = "l", lwd=3, col="black")
obj3 <- xyplot(pe ~ x, data, type = "l" , lwd=3, col="green")
obj4 <- xyplot(Aopt$w.best ~ x, data, type = "h", lwd=3.5,
              col="blue",ylim=c(0,30),ylab="")
trellis.par.set(axis.text=list(cex=1.6),ylab=list(col="blue"))
doubleYScale(obj1+obj2+obj3, obj4, add.ylab2 = TRUE,use.style=FALSE)
```

## Príloha C.

Výpočty v softvéri R pre A- a D-optimálne návrhy pri EMAX modeli.

```
# matica planu
F.emax<-Fx_dose(1:150,c(0.55,0.48,15),dose.model="emax")
w1<-od_REX(F.emax)$w.best      # aproximativny D-optimalny navrh
opt1<-optcrit(F.emax,w1*100)

X<-1:150
t1<-c(0.55,0.48,15)      # theta pre efektivitu
pe<-rep(0,150)
for ( i in 1:150) { pe[i]<-t1[1]+(X[i]*t1[2])/(X[i]+t1[3]) }

t2<-c(0,0.267,50)      # theta pre toxicitu
pt<-rep(0,150)
for ( i in 1:150) { pt[i]<-t2[1]+(X[i]*t2[2])/(X[i]+t2[3]) }

# pravdepodobnost, ze liek je neucinnny a ma vedlajsie ucinky
pf<-rep(0,150)
for ( i in 1:150) { pf[i]<-1-(1-pt[i])*pe[i] }

b<-c(100,10)      # 100 pacientov, pocet max. zlyhani=10
A<-matrix(1,nrow=2,ncol=150)
for (i in 1:150) { A[2,i]<-pf[i] }

Dopt<-od_AQUA(F.emax,b,A,t.max=30)
wD<-Dopt$w.best      # exaktny D-opt.navrh

# D-efektivita
effD<-Dopt$Phi.best/opt1

w2<-od_REX(F.emax,crit="A")$w.best      # aproximativny A-optimalny navrh
opt2<-optcrit(F.emax,w2*100,crit="A")

Aopt<-od_AQUA(F.emax,b,A,t.max=30,crit="A")
wA<-Aopt$w.best      # exaktny A-opt.navrh

# A-efektivita
effA<-Aopt$Phi.best/opt2

#library(lattice)
#library(latticeExtra)
#graficke znazornenie

x <- 1:150
data <- data.frame(x,pf,pt,pe,Dopt$w.best)
obj1 <- xyplot(pf ~ x, data, type = "l" , lwd=3, col="red",ylim=c(0,1),
  scale=list(y=list(tck=0.5)),xlab=list("dávka x",fontsize=25),
  ylab=list("pravdepodobnost",fontsize=25),
  ylab.right=list("počet pacientov",fontsize=25,col="blue"),
  par.settings=list(layout.widths=list(ylab.axis.padding=3)))
```

```

obj2 <- xyplot(pt ~ x, data, type = "l", lwd=3, col="black")
obj3 <- xyplot(pe ~ x, data, type = "l" , lwd=3, col="green")
obj4 <- xyplot(Dopt$w.best ~ x, data, type = "h", lwd=3.5,
               col="blue",ylim=c(0,30),ylab="")
trellis.par.set(axis.text=list(cex=1.6),ylab=list(col="blue"))
doubleYScale(obj1+obj2+obj3, obj4, add.ylab2 = TRUE,use.style=FALSE)

data <- data.frame(x,pf,pt,pe,Aopt$w.best)
obj1 <- xyplot(pf ~ x, data, type = "l" , lwd=3, col="red",ylim=c(0,1),
               scale=list(y=list(tck=0.5)),xlab=list("dávka x",fontsize=25),
               ylab=list("pravdepodobnosť",fontsize=25),
               ylab.right=list("počet pacientov",fontsize=25,col="blue"),
               par.settings=list(layout.widths=list(ylab.axis.padding=3)))
obj2 <- xyplot(pt ~ x, data, type = "l", lwd=3, col="black")
obj3 <- xyplot(pe ~ x, data, type = "l" , lwd=3, col="green")
obj4 <- xyplot(Aopt$w.best ~ x, data, type = "h", lwd=3.5,
               col="blue",ylim=c(0,30),ylab="")
trellis.par.set(axis.text=list(cex=1.6),ylab=list(col="blue"))
doubleYScale(obj1+obj2+obj3, obj4, add.ylab2 = TRUE,use.style=FALSE)

```



## Príloha D.

Výpočty v softvéri R pre A- a D-optimálne návrhy pri log-lineárnom modeli.

```
# matica planu
F.loglin<-Fx_dose(1:150,c(0,0.195,30),dose.model="loglin")

w1<-od_REX(F.loglin)$w.best # aproximativny D-opt.navrh
opt1<-optcrit(F.loglin,w1*100)

X<-1:150
t1<-c(0,0.195,30) #theta pre efektivitu
pe<-rep(0,150)
for ( i in 1:150) { pe[i]<-t1[1]+t1[2]*log(X[i]+t1[3]) }

t2<-c(0,0.02,0) #theta pre toxicitu
pt<-rep(0,150)
for ( i in 1:150) { pt[i]<-t2[1]+t2[2]*log(X[i]+t2[3]) }

# pravdepodobnost, ze liek je neucinnny a ma vedlajsie ucinky
pf<-rep(0,150)
for ( i in 1:150) { pf[i]<-1-(1-pt[i])*pe[i] }

b<-c(100,10) # 100 pacientov, pocet max. zlyhani=10
A<-matrix(1,nrow=2,ncol=150)
for (i in 1:150) { A[2,i]<-pf[i] }

Dopt<-od_AQUA(F.loglin,b,A,t.max=30)
wD<-Dopt$w.best # exaktny D-opt.navrh

# D-efektivita
effD<-Dopt$Phi.best/opt1

w2<-od_REX(F.loglin,crit="A")$w.best # aproximativny A-opt.navrh
opt2<-optcrit(F.loglin,w2*100,crit="A")

Aopt<-od_AQUA(F.loglin,b,A,t.max=30,crit="A")
wA<-Aopt$w.best # exaktny A-opt.navrh

# A-efektivita
effA<-Aopt$Phi.best/opt2

#library(lattice)
#library(latticeExtra)
#graficke znazornenie

x <- 1:150
data <- data.frame(x,pf,pt,pe,Dopt$w.best)
obj1 <- xyplot(pf ~ x, data, type = "l" , lwd=3, col="red",ylim=c(0,1),
  scale=list(y=list(tck=0.5)),xlab=list("dávka x",fontsize=25),
  ylab=list("pravdepodobnost",fontsize=25),
  ylab.right=list("počet pacientov",fontsize=25,col="blue"),
```

```

        par.settings=list(layout.widths=list(ylab.axis.padding=3)))
obj2 <- xyplot(pt ~ x, data, type = "l", lwd=3, col="black")
obj3 <- xyplot(pe ~ x, data, type = "l" , lwd=3, col="green")
obj4 <- xyplot(Dopt$w.best ~ x, data, type = "h", lwd=3.5,
              col="blue",ylim=c(0,30),ylab="")
trellis.par.set(axis.text=list(cex=1.6),ylab=list(col="blue"))
doubleYScale(obj1+obj2+obj3, obj4, add.ylab2 = TRUE,use.style=FALSE)

data <- data.frame(x,pf,pt,pe,Aopt$w.best)
obj1 <- xyplot(pf ~ x, data, type = "l" , lwd=3, col="red",ylim=c(0,1),
              scale=list(y=list(tck=0.5)),xlab=list("dávka x",fontsize=25),
              ylab=list("pravdepodobnosť",fontsize=25),
              ylab.right=list("počet pacientov",fontsize=25,col="blue"),
              par.settings=list(layout.widths=list(ylab.axis.padding=3)))
obj2 <- xyplot(pt ~ x, data, type = "l", lwd=3, col="black")
obj3 <- xyplot(pe ~ x, data, type = "l" , lwd=3, col="green")
obj4 <- xyplot(Aopt$w.best ~ x, data, type = "h", lwd=3.5,
              col="blue",ylim=c(0,30),ylab="")
trellis.par.set(axis.text=list(cex=1.6),ylab=list(col="blue"))
doubleYScale(obj1+obj2+obj3, obj4, add.ylab2 = TRUE,use.style=FALSE)

```