

# Software GramCharlier

*nástroj pro analýzu nejistot s využitím metody Gram-Charlier Expansion*

*Lukáš Novák (novak.l@fce.vutbr.cz), Drahomír Novák*

*Ústav stavební mechaniky, Fakulta stavební, Vysoké učení technické v Brně*

<https://www.fce.vutbr.cz/stm/novak.l/GCexp/GCexp.html>

## 1. Motivace a teoretický základ

Pro korektní reprezentaci reálných konstrukcí pomocí matematických modelů je nezbytné, aby bylo využito nejen moderních numerických algoritmů, jako například nelineární metody konečných prvků (NLMKP) často implementované ve statickém software, ale také uvažovat vstupní veličiny jako nejisté. V takovém případě je nezbytnou součástí procesu analýza nejistot. Analýza nejistot matematických modelů reprezentujících fyzikální systémy (např. stavební konstrukce) představuje důležitou část výzkumu posledních let s mnoha praktickými důsledky, které je nutné brát v potaz. Ačkoliv bylo v poslední dekádě vyvinuto mnoho efektivních metod pro propagaci nejistot vstupních parametrů matematickým modelem a následnou kvantifikaci nejistot zájmové veličiny (odezvy modelu), přesnost získaných výsledků je zcela zásadně ovlivněna předpokládaným rozdělením vstupních veličin.

Odhad rozdělení pravděpodobnosti vstupních veličin je typicky empirický na základě histogramu získaného z měření. Další možností je však kalibrace parametrů obecného pravděpodobnostního rozdělení tak, aby vystihoval zájmovou veličinu. Pro tvorbu unikátního rozdělení je nutné obecně znát nekonečně mnoho statistických momentů náhodné veličiny, což však není realistické pro inženýrské úlohy. Praktický a efektivní přístup nabízí zjednodušená metoda Gram-Charlier expansion, jež vyžaduje odhad prvních čtyř statistických momentů, na jejichž základě upraví tvar Gaussovského rozdělení tak, aby vyhovovalo dané veličině  $Y$ .

### I. Gram-Charlier Expansion

Hustota pravděpodobnosti  $f_Y$  je definována pomocí Gram-Charlier expansion [1] ve formě perturbace hustoty standardního Gaussovského rozdělení  $\varphi(y)$  jako:

$$f_Y(y) = \left[ 1 + \frac{\gamma_Y}{\sqrt{3!}} H_3(y) + \frac{\kappa_Y - 3}{\sqrt{4!}} H_4(y) \right] \varphi(y),$$

kde  $H_3$  a  $H_4$  představují Hermitovské polynomy třetího a čtvrtého řádu,  $\gamma_Y$  představuje koeficient šikmosti (třetí centrální moment) a  $\kappa_Y$  koeficient špičatosti (čtvrtý centrální moment). První dva statistické momenty jsou zohledněny při standardizaci veličiny  $Y$ , tak aby její střední hodnota  $\mu = 0$  a rozptyl  $\sigma^2 = 1$ .

Distribuční funkce odvozena na základě Gram-Charlier expansion je poté v následující formě:

$$F_Y(y) = \Phi(y) - \left[ \frac{\gamma_Y}{3\sqrt{2!}} H_2(y) + \frac{\kappa_Y - 3}{4\sqrt{3!}} H_3(y) \right] \varphi(y)$$

kde  $\Phi(y)$  představuje distribuční funkci standardního Gaussovského rozdělení. Odvození funkce a další teoretické detaily lze nalézt v publikaci autora softwarového nástroje [2].

## II. Odhad kvantilů

Kvantilová funkce neboli inverzní distribuční funkce  $F_Y^{-1}$  není bohužel dostupná v uzavřeném tvaru, a tudíž je nutné její hodnotu stanovit pomocí optimalizačních technik. Tento přístup však není typicky výpočetně náročný, jelikož se jedná o optimalizaci parametrů jednoduché funkce se skalární vstupní veličinou.

## III. Odhad statistických momentů a Bayesovská statistika

Přesnost pravděpodobnostního modelu s využitím Gram-Charlier expansion je přímo dána přesností odhadů statistických momentů zájmové veličiny. Software GramCharlier je navržen pro analýzu veličin, u nichž je možné získat velké množství měření a tudíž lze přímo využít numerický odhad všech momentů. V případě, že by byl software využit na základě nižšího počtu měření, je možné využít kombinaci naměřených dat a předpokládaného rozdělení z literatury s pomocí Bayesovské statistiky [3].

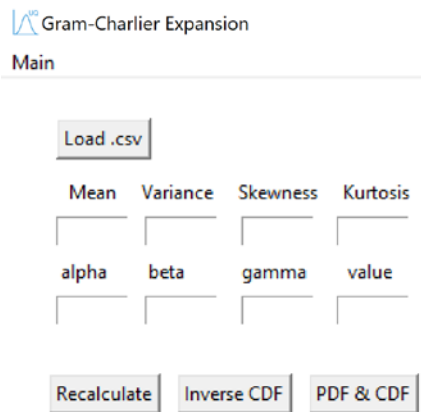
## 2. Uživatelský manuál

Na základě Gram-Charlier expansion byl vytvořen komplexní softwarový nástroj GramCharlier, jež je určen pro analýzu naměřených dat, odhad rozdělení pravděpodobnosti zájmové veličiny a jejího využití pro kvantifikaci nejistot včetně odhadu kvantilů potřebných pro polo-pravděpodobnostní analýzu.

Úvodní obrazovka softwaru GramCharlier je zobrazena na Obr. 1. Klasické rolovací menu „Main“ obsahuje odkaz na aktualizovaný manuál, možnost restartovat výpočet a také ukončit program. Na hlavní obrazovce je poté tlačítko Load .csv, jež umožní uživateli načíst změřená data ve formátu .csv. Po načtení dat, jsou automaticky odhadnuty první čtyři statistické momenty (mean, variance, skewness, kurtosis) s využitím standardních přístupů matematické statistiky. Na základě těchto momentů je vytvořen pravděpodobnostní model zájmové veličiny ve formě Gram-Charlier expansion. Následně jsou vykresleny funkce hustoty a distribuční funkce společně s histogramy empirického rozdělení pro vizuální kontrolu (viz Obr. 2). Momenty lze také vepsat či upravit manuálně, což je vhodné využít při Bayesovské kalibraci pomocí softwaru třetích stran. Po manuální úpravě hodnot, je vždy nutné opětovně sestavit Gram-Charlier expansion pomocí tlačítka recalculate.

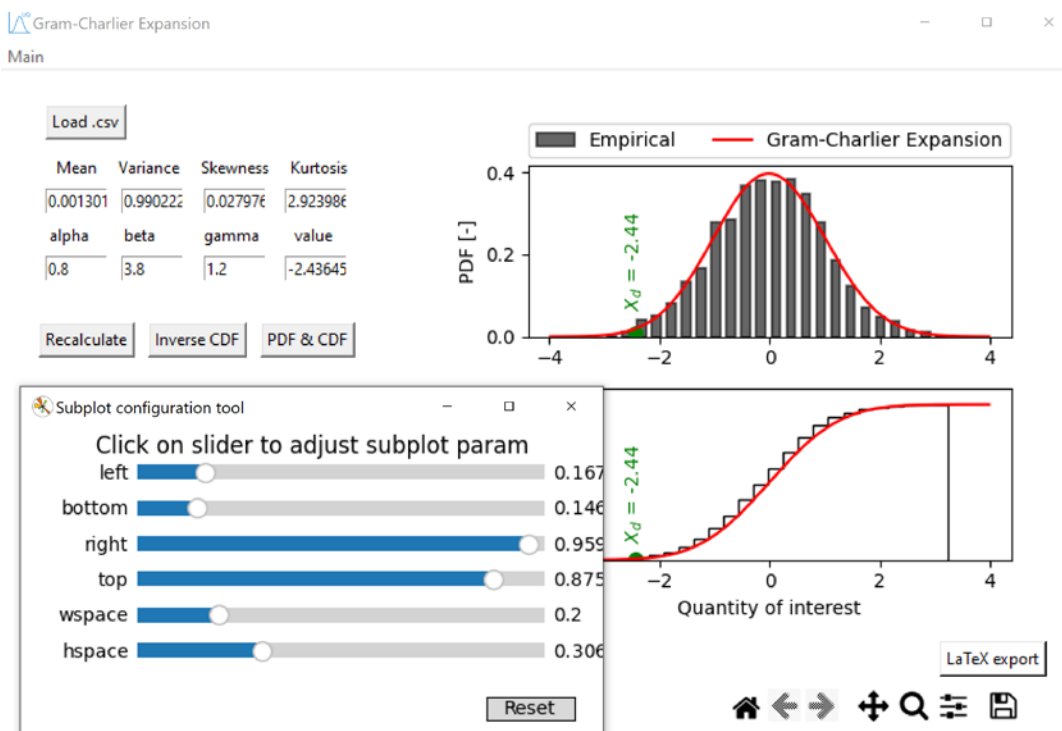
Druhý řádek obsahuje vstupní pole pro zadání hodnot nutných při polo-pravděpodobnostním přístupu pro odhad návrhových kvantilů (index spolehlivosti Beta, citlivostní součinitel Alpha a dílčí součinitel modelových nejistot Gamma). Při zadání všech tří

polí je odhadnut požadovaný kvantil pomocí optimalizačního algoritmu. Hodnota návrhového kvantilu je zobrazena v poli „value“ a zobrazena na pravé straně programu (viz Obr. 2).



Obr. 1 Úvodní obrazovka softwaru GramCharlier

Poslední řádek obsahuje tlačítka pro post-processing. Konkrétně jde o stanovení hodnot hustoty pravděpodobnosti (PDF) a distribuční funkce (CDF) pro zadané naměřené hodnoty v souboru .csv. Pro další statistické zpracování je často nutné provést statistické vzorkování z identifikovaného rozdělení pravděpodobnosti, což lze provést pomocí tlačítka „Inverse CDF“, jež umožňuje načíst .csv soubor obsahující hodnoty pravděpodobnosti odpovídající pozici požadovaných statistických vzorků. U obou funkcí je tedy nejprve načten soubor .csv, následně je provedena požadovaná akce a nakonec program vytvoří nový soubor ve formátu .csv obsahující zadané vstupní hodnoty spolu se stanovenými výsledky.

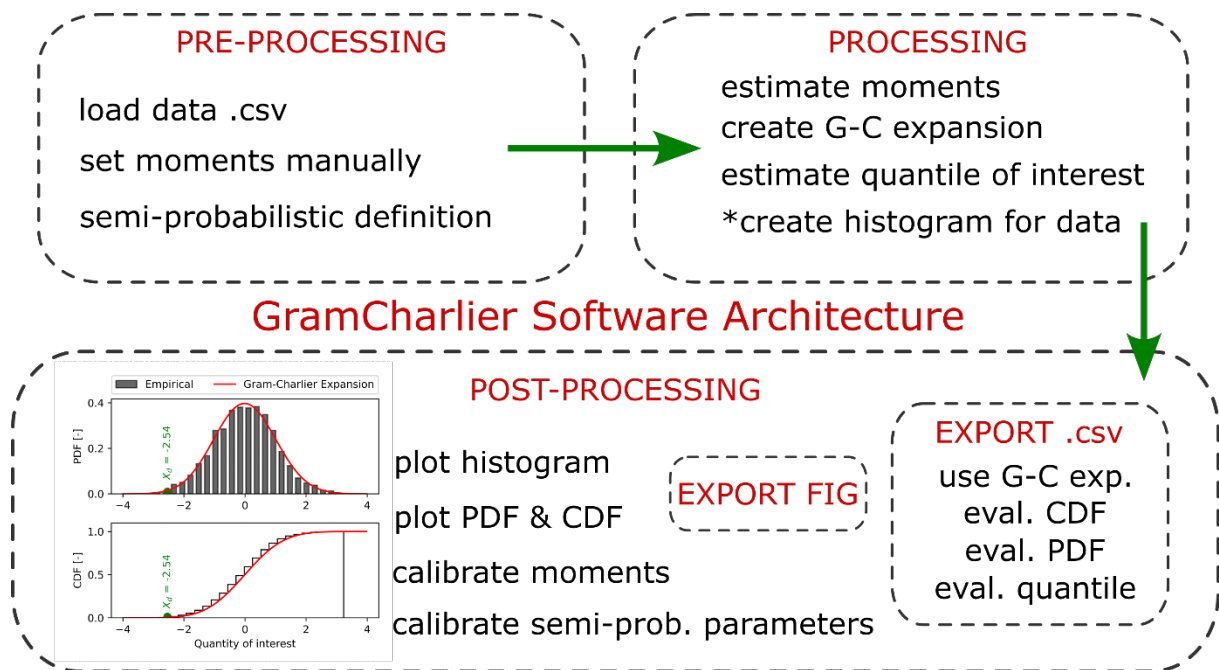


Obr. 2 Kompletní obrazovka softwaru GramCharlier.

Poslední funkcí softwaru GramCharlier je tvorba detailních grafů zobrazujících empirické rozdělení (histogram) společně s pravděpodobnostním modelem Gram-Charlier expansion. Navíc lze také zobrazit odhadovaný návrhový kvantil s využitím polo-pravděpodobnostního přístupu. Pod samotným grafem je poté nástrojová lišta obsahující hlavní funkce pro úpravu zobrazení, detailní nastavení (viz Obr. 2) lze nalézt pod tlačítkem settings. Tlačítko uložit umožňuje export vytvořeného grafu. Za zmínku jistě stojí také tlačítko „LaTeX export“ umožňující vysoce kvalitní renderování grafu s využitím systému LaTeX a produkci vysoce kvalitních .PDF souborů, využitelných při tvorbě reportů či vědeckých publikací.

### 3. Analýza funkčních požadavků a koncepce software GramCharlier

Na základě důkladné analýzy dostupné literatury a požadavků potenciačních uživatelů, byla vytvořena následující koncepce zobrazena na Obr. 3 ve formě blokového diagramu. Prioritou byla uživatelská přívětivost a možnost použít software bez hlubších znalostí teorie pravděpodobnosti a matematické statistiky. Z toho důvodu byl vytvořen zcela automatický nástroj pro tvorbu Gram-Charlier expansion, který vyžaduje od uživatele pouze načtení .csv souboru obsahujícího měření. Celý proces se dá rozdělit do tří imaginárních částí: pre-processing, processing a post-processing.



Obr. 3 Základní koncepce softwaru GramCharlier

#### I. Pre-processing

Pre-processing neboli přípravná fáze procesu, je zaměřena na získání vstupních dat od uživatele. Konkrétně lze načíst měření ze souboru ve formátu .csv, nastavit hodnotu prvních čtyř statistických momentů manuálně, a v neposlední řadě stanovit požadované koeficienty pro odhad návrhového kvantilu s využitím polo-pravděpodobnostního přístupu.

## II. Processing

Jádrem programu je plně automatická část processing, která zajišťuje odhad a zpracování statistických momentů, tvorbu histogramu v případě načtených měření, sestavení kompletního pravděpodobnostního modelu zájmové veličiny s využitím Gram-Charlier expansion a v neposlední řadě stanovení návrhového kvantilu.

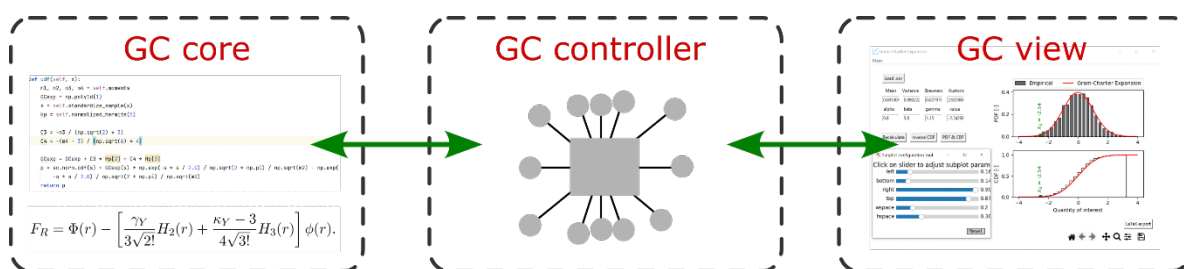
## III. Post-processing

V rámci poslední části post-processing, jsou veškeré stanovené údaje zobrazeny ve formě dvou grafů: hustota pravděpodobnosti a distribuční funkce. Oba grafy mohou být dále doplněny o empirické rozdělení (histogram) a označení návrhového kvantilu. Dále je možné v rámci post-processingu kalibrovat zadané statistické momenty či údaje pro polo-pravděpodobnostní výpočet. Vytvořené grafy lze následně exportovat v různých grafických formátech, včetně PDF vytvořeného pomocí systému LaTeX. Navíc lze využít vytvořený pravděpodobnostní model využít pro stanovení PDF a CDF daných měření (ve formátu .csv), či pro statistické vzorkování pro zadané pravděpodobnosti výskytu (opět ve formátu .csv).

## 4. Technická dokumentace a zdrojový kód

Z technického pohledu je software vytvořen s využitím zásad objektově orientovaného programování ze 3 tříd: GCcore, GCcontroller a GCview (viz Obr. 4). Třída GCcore obsahuje všechny matematické funkce důležité pro tvorbu a zpracování pravděpodobnostního modelu Gram-Charlier expansion. Třída GCView obsahuje veškeré prvky potřebné pro grafické uživatelské rozhraní. Třída GCcontroller zprostředkovává komunikaci mezi všemi třídami. Takto navrženy software umožňuje efektivní správu kódu a jednoduché rozšíření o další funkce. Samotný softwarový nástroj je vytvořen v jazyce Python s grafickým modulem Tkinter.

Kompletní zdrojový kód aplikace je veřejně přístupný na cloudovém uložišti GitHub: <https://github.com/NovakLBUT/GramCharlier>



Obr. 4 Třídy tvořící zdrojový kód softwaru GramCharlier.

## 5. Validace a závěrečné poznámky

Vytvořený software GramCharlier představuje efektivní nástroj pro analýzu nejistot výpočetně náročných matematických modelů a je tedy vhodným doplňkem praktického návrhu a posouzení složitých konstrukcí. Využití GramCharlier je intuitivní a neklade vysoké nároky na koncové uživatele. Teoretické detaily jsou k dispozici v publikovaném vědeckém článku autora softwaru [2]. Validace a verifikace softwarového nástroje byla provedena autorským týmem přepočtem několika teoretických i praktických úloh [3].

*Vývoj softwarového nástroje GramCharlier byl finančně podpořen projektem Technologické agentury České republiky (TAČR) evidovaným pod číslem TM04000012.*

## Reference

- [1] D. E. Barton, K. E. Dennis, The conditions under which Gram-Charlier and Edgeworth curves are positive definite and unimodal, *Biometrika* 39 (3/4) (1952) 425–427. doi:10.2307/2334037.
- [2] L. Novák, On Distribution-Based Global Sensitivity Analysis by Polynomial Chaos Expansion. *Computers & Structures*, 267 (2022), ISSN: 0045-7949.
- [3] L. Novák, D. Novák, Semi-Probabilistic Assessment of Concrete Bridge Using Polynomial Chaos and Gram-Charlier Expansions, In *ENGINEERING MECHANICS 2022 PROCEEDINGS, VOL 27/28 289-292*, Praha: Institute of Theoretical and Applied Mechanics of the Czech Academy of Sciences, (2022), ISBN: 978-80-214-5896-3.