Simulationstools für B-SDC

Übung 3 **(Mangold)** **Wintersemester 2025/26**

Kapitel 3: Varianzreduktion

Hinweis: Falls nicht anders angegeben, verwenden Sie NumPy/SciPy. RNG: z. B. PCG64, Basis-Seed 42 (ggf. Substreams).

Aufgabe 10 Antithetik vs. naiv

Vergleichen Sie Standardfehler (SE) naiv vs. antithetisch für $\mu = \mathbb{E}[\sqrt{X}]$ mit $X \sim \mathcal{U}(0, 1)$.

a) Schreiben Sie eine Funktion, welche die Schätzungen und SEs für beide Szenarien berechnet:
   - **Naiv:** Ziehen Sie $X_i \sim \mathcal{U}(0, 1)$, $g(X_i) = \sqrt{X_i}$, $\hat{\mu} = \frac{1}{N} \sum g(X_i)$.
   - **Antithetik:** Nutzen Sie $N/2$ Ziehungen U_i. Bilden Sie Paare $(U_i, 1 - U_i)$ und die zugehörigen g-Werte. Mitteln Sie die $N/2$ paarweisen Durchschnitte $\frac{1}{2}(g(U_i) + g(1 - U_i))$.

b) Erstellen Sie eine Tabelle für $N \in \{10^3, 10^4, 10^5\}$. Vergleichen Sie für beide Methoden $\hat{\mu}$, den geschätzten SE und das 95%-CI.

c) Fügen Sie eine Spalte für den **Gewinnfaktor** $\left(\frac{\mathrm{SE}_{\mathrm{naiv}}}{\mathrm{SE}_{\mathrm{anti}}}\right)$ hinzu und interpretieren Sie das Ergebnis: Warum funktioniert Antithetik hier so gut?

Aufgabe 11 Control-Variate (CV)

Wir schätzen weiterhin $\mu = \mathbb{E}[g(X)]$ mit $g(X) = \sqrt{X}$ für $X \sim \mathcal{U}(0, 1)$. Verwenden Sie nun $h(X) = X$ als Control Variate. Wir wissen analytisch, dass $\mathbb{E}[h(X)] = \mu_h = 0{,}5$.

a) Schätzen Sie den optimalen Koeffizienten $\hat{\beta}^\star = \widehat{\mathrm{Cov}}(g, h)/\widehat{\mathrm{Var}}(h)$ empirisch aus $N = 50\,000$ Samples.

b) Berechnen Sie den CV-Schätzer $\hat{\mu}_{\mathrm{CV}} = \bar{g} - \hat{\beta}(\bar{h} - 0{,}5)$.

c) Vergleichen Sie den $\widehat{\mathrm{SE}}(\hat{\mu}_{\mathrm{CV}})$ mit dem $\widehat{\mathrm{SE}}(\hat{\mu}_{\mathrm{naiv}})$ für $N \in \{5\,000, 50\,000\}$. (*Schätzen Sie den SE durch Replikation.*)

d) Weisen Sie die SE-Reduktion in % aus und begründen Sie (z.B. anhand von $\hat{\rho}_{gh}$), warum $h(X) = X$ hierfür geeignet ist.

Aufgabe 12 LHS (2D) vs. i.i.d.

Zeigen Sie, dass Latin Hypercube Sampling (LHS) bei einem glatten Integranden die Schätzvarianz gegenüber i.i.d.-Sampling reduziert. Geschätzt werden soll $\mu = \mathbb{E}[g(Z_1, Z_2)]$ mit

$$g(Z_1, Z_2) = \exp\Big(- 0{,}5\,(Z_1^2 + Z_2^2)\Big),$$

wobei Z_1, Z_2 unabhängige $\mathcal{N}(0,1)$ Variablen sind. (Sie werden via $\Phi^{-1}(U)$ aus $U \sim \mathcal{U}[0,1]^2$ erzeugt).

a) **Baseline (i.i.d.):** Ziehen Sie N Punkte $(Z_{1,i}, Z_{2,i})$ i.i.d. aus $\mathcal{N}(0, I_2)$. Berechnen Sie $\hat{\mu}_{\text{naiv}}$ und $\widehat{\text{SE}}_{\text{naiv}}$.

b) **LHS (2D):** Erzeugen Sie ein $N \times 2$ LHS-Design auf $[0,1]^2$. Transformieren Sie beide Spalten via Φ^{-1} (PPF) zu $(Z_{1,i}, Z_{2,i})$.

c) **SE-Schätzung für LHS:** Da die LHS-Samples nicht i.i.d. sind verwenden Sie die **Replikations-Methode**:

- Führen Sie $R = 50$ unabhängige LHS-Läufe mit je $m = N/R$ Samples durch (z.B. für $N = 50\,000 \to R = 50, m = 1000$).
- Berechnen Sie $\hat{\mu}_{\text{LHS}}$ als Mittelwert der R Läufe und $\widehat{\text{SE}}_{\text{LHS}}$ als Standardfehler dieser R Mittelwerte.

d) **Vergleich:** Erstellen Sie eine Tabelle für $N_{\text{total}} \in \{10^4, 5 \cdot 10^4, 10^5\}$. Vergleichen Sie Naiv vs. LHS (mit Replikation) bzgl.: $\hat{\mu}$, SE, CPU-Zeit und der Effizienz-Metrik Effizienz $= \text{SE}^2 \times$ Zeit. Interpretieren Sie das Ergebnis.

e) **Visualisierung:** Erstellen Sie einen Scatterplot ($N = 500$) der (Z_1, Z_2)-Punkte für i.i.d. vs. LHS, um die unterschiedliche Raumabdeckung zu zeigen.

Aufgabe 13 CRN und 2x1-Faktorielles Design

Wir vergleichen zwei Policies, A ("Standard") und B ("Neu"), deren Kosten $g(Z)$ von einem unsicheren Marktfaktor $Z \sim \mathcal{N}(\mu, 1)$ abhängen.
Die Kostenfunktionen seien (stark vereinfacht):

- **Policy A (Standard):** $g_A(Z) = 5 + \exp(Z)$
- **Policy B (Neu):** $g_B(Z) = 7 + 0.9 \cdot \exp(Z)$

Wir wollen die **Kostenersparnis** $\delta = \mathbb{E}[g_A(Z)] - \mathbb{E}[g_B(Z)]$ schätzen.

a) **CRN-Vergleich (für $\mu = 0$):** Schätzen Sie $\hat{\delta}$ für $Z \sim \mathcal{N}(0,1)$ mit $N = 50\,000$.

   (i) Einmal **ohne CRN** (unabhängige Samples für A und B).

(ii) Einmal **mit CRN** (dieselben Samples Z_i für A und B).

Erstellen Sie eine Tabelle: Vergleichen Sie $\hat{\delta}$, $\widehat{\mathrm{SE}}(\hat{\delta})$ und die Korrelation $\hat{\rho}(g_A, g_B)$ für (i) und (ii). Quantifizieren Sie den Gewinnfaktor ($\mathrm{SE}_{\text{ohne}}/\mathrm{SE}_{\text{mit}}$).

b) **2x1-Faktorielles Design:** Wie ändert sich die Ersparnis δ, wenn sich der Markt ändert?

- Schätzen Sie $\hat{\delta}_0$ für **Faktor** $\mu = 0$ (mit CRN, $N = 50\,000$).
- Schätzen Sie $\hat{\delta}_1$ für **Faktor** $\mu = 1$ (mit CRN, $N = 50\,000$).

Berechnen Sie den **Haupteffekt** von μ auf die Ersparnis: $\text{Effekt}_\mu = \hat{\delta}_1 - \hat{\delta}_0$. Interpretieren Sie das Ergebnis in 1-2 Sätzen. (z.B. "Wenn der Markt (mu) steigt, wird Policy B [besser/schlechter] im Vergleich zu A.")