# week8: Monte Carlo Simulation, Teil2

Erinnern wir uns an die Monte Carlo Sachen aus dem `week7.pdf`: Ziel war es, das Integral oder den Erwartungswert

$$I \;:=\; \int_{-\infty}^{+\infty} F(x)\, p(x)\, dx \;=\; \mathsf{E}[\,F\,] \tag{1}$$

numerisch zu berechnen. Dazu hatten wir $N$ $p(x)$-verteilte Zufallszahlen

$$x_1 \,,\; x_2 \,,\; \cdots \;,\; x_N \tag{2}$$

generiert und dann die Monte Carlo Summe

$$S_N(F) \;:=\; \frac{1}{N} \sum_{i=1}^{N} F(x_i) \tag{3}$$

gebildet. Für grosse $N$ konvergiert diese Summe dann gegen den Erwartungswert von $F$,

$$\lim_{N\to\infty} S_N(F) \;=\; \mathsf{E}[\,F\,] \;=\; \int_{-\infty}^{+\infty} F(x)\, p(x)\, dx \tag{4}$$

Genauer hatten wir in (8) im week7.pdf die folgende Abschätzung angegeben, die wir hier jetzt also beweisen wollen:

$$\mathsf{Prob}\Big[\, \big|\, S_N(F) - \mathsf{E}[F] \,\big| \;\geq\; \varepsilon \,\Big] \;\leq\; \frac{\mathsf{V}[F]}{N\varepsilon^2} \tag{5}$$

Zum Beweis von (5) müssen wir den Erwartungswert und die Varianz der Monte Carlo Summe $S_N(F)$ berechnen.

Das $S_N(F)$ hängt von $N$ Zufallszahlen ab, nicht nur von einer. Wenn man also einen Erwartungswert für das $S_N(F)$ hinschreiben möchte, muss man jede einzelne Zufallszahl $x_i$ gegen ihre Dichte $p(x_i)$ integrieren:

$$\mathsf{E}\big[\, S_N(F) \,\big] \;=\; \int_{\mathbb{R}^N} S_N(F)\, p(x_1)\, p(x_2) \cdots p(x_N)\, dx_1\, dx_2 \cdots dx_N \tag{6}$$

Wegen

$$\int_{\mathbb{R}} p(x_i)\, dx_i \;=\; 1 \tag{7}$$

lässt sich der Ausdruck (6) aber sehr stark vereinfachen: Wir bekommen

$$
\begin{aligned}
\mathsf{E}\big[\, S_N(F) \,\big] \;&=\; \int_{\mathbb{R}^N} S_N(F)\, p(x_1)\, p(x_2) \cdots p(x_N)\, dx_1\, dx_2 \cdots dx_N \\
&=\; \int_{\mathbb{R}^N} \frac{1}{N} \sum_{i=1}^{N} F(x_i)\;\; p(x_1)\, p(x_2) \cdots p(x_N)\, dx_1\, dx_2 \cdots dx_N \\
&=\; \frac{1}{N} \sum_{i=1}^{N} \int_{\mathbb{R}^N} F(x_i)\;\; p(x_1)\, p(x_2) \cdots p(x_N)\, dx_1\, dx_2 \cdots dx_N \\
&=\; \frac{1}{N} \sum_{i=1}^{N} \underbrace{\int_{\mathbb{R}} F(x_i)\, p(x_i)\, dx_i}_{=\,\mathsf{E}[F]} \prod_{\substack{j=1 \\ j\neq i}}^{n} \underbrace{\int_{\mathbb{R}} p(x_j)\, dx_j}_{=\,1} \\
&=\; \frac{1}{N} \sum_{i=1}^{N} \mathsf{E}[\, F \,] \;=\; \mathsf{E}[\, F \,] \qquad (8)
\end{aligned}
$$

Als nächstes wollen wir die Varianz von dem $S_N(F)$ berechnen. Um die Rechnung möglichst übersichtlich zu gestalten, definieren wir zunächst die Kovarianz von zwei zufälligen Grössen $X$ und $Y$:

**Definition 8.1 (Kovarianz):** Die Kovarianz von zwei zufälligen Grössen $X$ und $Y$ ist definiert durch

$$
\mathsf{Cov}[X, Y] \;:=\; \mathsf{E}\Big[\, \big(\, X - \mathsf{E}[X] \,\big)\, \big(\, Y - \mathsf{E}[Y] \,\big) \,\Big] \;.
$$

**Lemma 8.1:** Die Kovarianz hat die folgenden elementaren Eigenschaften:

**a)** Es gilt

$$
\mathsf{Cov}[X, Y] \;=\; \mathsf{E}[\, XY \,] \;-\; \mathsf{E}[X]\, \mathsf{E}[Y]
$$

und

$$
\mathsf{Cov}[X, X] \;=\; \mathsf{V}[X]
$$

**b)** Die Kovarianz ist bilinear, es gilt

$$
\begin{aligned}
\mathsf{Cov}[\, \lambda_1 X_1 + \lambda_2 X_2 \,,\, Y \,] \;&=\; \lambda_1\, \mathsf{Cov}[X_1, Y] \;+\; \lambda_2\, \mathsf{Cov}[X_2, Y] \\
\mathsf{Cov}[\, X \,,\, \mu_1 Y_1 + \mu_2 Y_2 \,] \;&=\; \mu_1\, \mathsf{Cov}[X, Y_1] \;+\; \mu_2\, \mathsf{Cov}[X, Y_2]
\end{aligned}
$$

**Beweis:** Übungsblatt 8. ■

**Beispiel:** Es seien $x_i$ und $x_j$ die $i$-te und die $j$-te Zufallszahl in der Monte Carlo Summe $S_N(F)$ und $i \neq j$. Wir betrachten die Zufallsgrössen

$$
\begin{aligned}
X \;&:=\; F(x_i) \\
Y \;&:=\; F(x_j)
\end{aligned}
$$

Dann gilt

$$\begin{aligned}
\mathsf{E}[\,XY\,] &= \int_{\mathbb{R}^2} F(x_i)\,F(x_j)\,p(x_i)\,p(x_j)\,dx_i\,dx_j \\
&= \int_{\mathbb{R}} F(x_i)\,p(x_i)\,dx_i\,\int_{\mathbb{R}} F(x_j)\,p(x_j)\,dx_j \\
&= \mathsf{E}[\,F\,]\,\mathsf{E}[\,F\,] \;=\; \big(\mathsf{E}[\,F\,]\big)^2
\end{aligned}$$

und damit

$$\mathsf{Cov}[X,Y] \;=\; \mathsf{E}[\,XY\,] - \mathsf{E}[X]\,\mathsf{E}[Y] \;=\; \mathsf{E}[\,F\,]\,\mathsf{E}[\,F\,] - \mathsf{E}[\,F\,]\,\mathsf{E}[\,F\,] \;=\; 0$$

Für $i = j$ bekommt man die Varianz von $F$,

$$\begin{aligned}
\mathsf{Cov}\big[\,F(x_i)\,,\,F(x_i)\,\big] \;=\; \mathsf{V}\big[\,F(x_i)\,\big] &= \int_{\mathbb{R}} \big(\,F(x_i) - \mathsf{E}[F]\,\big)^2\,p(x_i)\,dx_i \\
&= \int_{\mathbb{R}} \big(\,F(x) - \mathsf{E}[F]\,\big)^2\,p(x)\,dx \;=\; \mathsf{V}[F] \;.
\end{aligned}$$

**Definition 8.2:** Zwei zufällige Grössen $X$ und $Y$ heissen unabhängig, falls ihre Kovarianz 0 ist,

$$\mathsf{Cov}[X,Y] \;=\; 0 \;.$$

Gemäss Definition 8.2 sind $F(x_i)$ und $F(x_j)$ dann also unabhängige Grössen. Mit diesen Vorbereitungen können wir jetzt die Varianz der Monte Carlo Summe $S_N(F)$ berechnen:

$$\begin{aligned}
\mathsf{V}\big[\,S_N(F)\,\big] &= \mathsf{V}\Big[\,\frac{1}{N}\sum_{i=1}^{N} F(x_i)\,\Big] \\
&= \mathsf{Cov}\Big[\,\frac{1}{N}\sum_{i=1}^{N} F(x_i)\,,\;\frac{1}{N}\sum_{j=1}^{N} F(x_j)\,\Big] \\
&\overset{\substack{\text{Cov ist}\\\text{bilinear}}}{=} \frac{1}{N^2}\sum_{i,j=1}^{N} \underbrace{\mathsf{Cov}\big[\,F(x_i)\;,\;F(x_j)\,\big]}_{=\,0 \text{ falls } i \neq j} \\
&= \frac{1}{N^2}\sum_{i=1}^{N} \mathsf{Cov}\big[\,F(x_i)\,,\,F(x_i)\,\big] \\
&= \frac{1}{N^2}\sum_{i=1}^{N} \mathsf{V}[\,F\,] \;=\; \frac{1}{N^2}\,N\,\mathsf{V}[\,F\,] \;=\; \frac{\mathsf{V}[\,F\,]}{N}
\end{aligned} \tag{9}$$

Das Wesentliche ist, dass da ein $1/N$ übrig bleibt, eigentlich hat man $N^2$ Terme in der Summe, aber da man eine Summe von unabhängigen Zufallsgrössen hat, sind alle 'off-diagonalen' Terme 0, das sind $N^2 - N$ Stück, und nur die $N$ Terme für $i = j$ liefern einen Beitrag, so hat man dann insgesamt nur $N\mathsf{V}[F]/N^2 = \mathsf{V}[F]/N$ und das geht nach 0 für $N \to \infty$. Wenn die Varianz 0 wird, meint das aber gerade, dass die Zufallsgrösse selber, hier also die Monte Carlo Summe, gegen ihren Erwartungswert konvergieren muss, und der war ja $\mathsf{E}[F] = \int_{\mathbb{R}} F(x)p(x)dx$ wie wir in (8) bereits berechnet hatten. Bevor wir das mit der Tschebychev'schen Ungleichung gleich noch etwas quantitativer formulieren, halten wir die Resultate für den Erwartungswert und die Varianz vielleicht in einem separatem Theorem fest:

**Theorem 8.1:** Es seien $x_1, \cdots, x_N$ $p(x)$-verteilte Zufallszahlen und $F = F(x) : \mathbb{R} \to \mathbb{R}$ eine beliebige Funktion. Wir bilden die Monte Carlo Summe

$$S_N(F) \;:=\; \frac{1}{N} \sum_{i=1}^{N} F(x_i)$$

Dann gilt:

$$\mathsf{E}\big[\, S_N(F) \,\big] \;=\; \mathsf{E}[F] \tag{10}$$

$$\mathsf{V}\big[\, S_N(F) \,\big] \;=\; \frac{\mathsf{V}[F]}{N} \tag{11}$$

mit dem Erwartungswert und der Varianz von $F$ geben durch

$$\begin{aligned} \mathsf{E}[F] &= \textstyle\int_{\mathbb{R}} F(x)\, p(x)\, dx \\ \mathsf{V}[F] &= \textstyle\int_{\mathbb{R}} \big( F(x) - \mathsf{E}[F] \big)^2\, p(x)\, dx \;\;. \end{aligned}$$

Um den Beweis der Abschätzung (5) abzuschliessen, brauchen wir jetzt nur noch die Tschebychev'sche Ungleichung, das ist das folgende allgemeine Theorem 8.2 weiter unten. Die Zufallsgrösse dort, das $X$, wäre in unserem Fall dann die Monte Carlo Summe,

$$X \;=\; S_N(F) \;:=\; \frac{1}{N} \sum_{i=1}^{N} F(x_i) \;=\; X(x_1, \cdots, x_N)$$

und die Wahrscheinlichkeitsdichte für die $N$ Zufallszahlen wäre dann also ein

$$p_N(x_1, \cdots, x_N) \;:=\; p(x_1)\, p(x_2)\; \cdots\; p(x_N) \;=\; \prod_{i=1}^{N} p(x_i) \tag{12}$$

**Theorem 8.2 (Tschebychev'sche Ungleichung):** Es seien $x_1, ..., x_N$ Zufallszahlen mit W'keitsdichte $p_N(x_1, ..., x_N)$ und $X = X(x_1, ..., x_N)$ sei eine beliebige zufällige Grösse. Dann gilt für jedes $\varepsilon > 0$:

$$\mathsf{Prob}\Big[\, \big|\, X - \mathsf{E}[X] \,\big| \geq \varepsilon \,\Big] \;\leq\; \frac{\mathsf{V}[X]}{\varepsilon^2} \;\;. \tag{13}$$

**Beweis:** Wir benutzen die folgenden Abkürzungen:

$$\begin{aligned} x &:= (x_1, \cdots, x_N) \\ d^N x &:= dx_1 \cdots dx_N \end{aligned}$$

Dann können wir die Varianz $\mathsf{V}[X]$ schreiben als

$$\mathsf{V}[X] \;=\; \textstyle\int_{\mathbb{R}^N} \big[\, X(x) \;-\; \mathsf{E}[X] \,\big]^2\, p_N(x)\, d^N x$$

Wir definieren die Mengen $A_\varepsilon, B_\varepsilon \subset \mathbb{R}^N$ durch

$$A_\varepsilon \;:=\; \left\{ x \in \mathbb{R}^N \;\middle|\; \big| X(x) - \mathsf{E}[X] \big| \;\geq\; \varepsilon \right\}$$
$$B_\varepsilon \;:=\; \left\{ x \in \mathbb{R}^N \;\middle|\; \big| X(x) - \mathsf{E}[X] \big| \;<\; \varepsilon \right\}$$

Dann gilt

$$\mathbb{R}^N \;=\; A_\varepsilon \,\cup\, B_\varepsilon$$

und wir können schreiben

$$\begin{aligned} \mathsf{V}[X] \;&=\; \int_{\mathbb{R}^N} \big[ X(x) \,-\, \mathsf{E}[X] \big]^2 \, p_N(x)\, d^N x \\ &=\; \int_{A_\varepsilon} \big[ X(x) \,-\, \mathsf{E}[X] \big]^2 \, p_N(x)\, d^N x \;+\; \int_{B_\varepsilon} \big[ X(x) \,-\, \mathsf{E}[X] \big]^2 \, p_N(x)\, d^N x \end{aligned}$$

Auf $A_\varepsilon$ ist

$$\big| X(x) - \mathsf{E}[X] \big| \;\geq\; \varepsilon \qquad \Leftrightarrow \qquad \big| X(x) - \mathsf{E}[X] \big|^2 \;\geq\; \varepsilon^2$$

Also können wir abschätzen:

$$\begin{aligned} \mathsf{V}[X] \;&=\; \int_{A_\varepsilon} \big[ X(x) \,-\, \mathsf{E}[X] \big]^2 \, p_N(x)\, d^N x \;+\; \int_{B_\varepsilon} \big[ X(x) \,-\, \mathsf{E}[X] \big]^2 \, p_N(x)\, d^N x \\ &\geq\; \int_{A_\varepsilon} \varepsilon^2 \, p_N(x)\, d^N x \;+\; \underbrace{\int_{B_\varepsilon} \big[ X(x) \,-\, \mathsf{E}[X] \big]^2 \, p_N(x)\, d^N x}_{\geq 0} \\ &\geq\; \int_{A_\varepsilon} \varepsilon^2 \, p_N(x)\, d^N x \;+\; 0 \\ &=\; \varepsilon^2 \int_{A_\varepsilon} p_N(x)\, d^N x \\ &=\; \varepsilon^2 \cdot \mathsf{Prob}[\, x \in A_\varepsilon \,] \;=\; \varepsilon^2 \cdot \mathsf{Prob}\Big[ \big| X - \mathsf{E}[X] \big| \geq \varepsilon \Big] \end{aligned}$$

also

$$\mathsf{V}[X] \;\geq\; \varepsilon^2 \cdot \mathsf{Prob}\Big[ \big| X - \mathsf{E}[X] \big| \geq \varepsilon \Big]$$

Wenn wir das durch $\varepsilon^2$ teilen, bekommen wir dann die Tschebychev'sche Ungleichung. ■

**Folgerung:** Wir wählen jetzt also

$$X \;:=\; S_N(F) \;=\; \frac{1}{N} \sum_{i=1}^{N} F(x_i)$$

in der Tschebychev'schen Ungleichung (13) und bekommen dann mit (10) und (11)

$$\mathsf{Prob}\Big[ \big| S_N(F) - \mathsf{E}[F] \big| \;\geq\; \varepsilon \Big] \;\leq\; \frac{\mathsf{V}[F]}{N\varepsilon^2}$$

und das war genau die Abschätzung (8) aus dem week7.pdf, die wir beweisen wollten. ■