Distributions

- Binomial distr

$$b_{n,p} = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{(n-k)}$$

$p \in [0,1]$ $n \in \mathbb{N}$

$b_{n,p} : \mathcal{B}(\mathbb{R}) \to [0,\infty]$

- Geometric distr

$$geom_p = \sum_{n=0}^{\infty} p(1-p)^n$$

$p \in (0,1]$

$geom_p : \mathcal{B}(\mathbb{R}) \to [0,\infty]$

- Poisson distr

$$Poi_\lambda = \sum_{n=0}^{\infty} \frac{e^{-\lambda}\lambda^n}{n!}$$

$\lambda \in (0,\infty)$

$Poi_\lambda : \mathcal{B}(\mathbb{R}) \to [0,\infty]$

↳ approx. for Binomial distr with $\lambda \approx np$ and n large, p small

- Cont. Uniform distr

$$U_A(B) = \frac{\lambda_{\mathbb{R}^d}(B \cap A)}{\lambda_{\mathbb{R}^d}(A)}$$

$d \in \mathbb{N}$; $A, B \in \mathcal{B}(\mathbb{R}^d)$

$0 < \lambda_{\mathbb{R}^d}(A) < \infty$

- Exponential distr

$$exp_\lambda(B) = \int_B \lambda e^{-\lambda x} \cdot \mathbb{1}^{\mathbb{R}}_{(0,\infty)}(x)\, dx$$

$\lambda \in (0,\infty)$

$$F(x) = 1 - e^{-\lambda x}$$

- Cauchy distr

$$Cau_{\mu,\lambda}(B) = \int_B \frac{dx}{\pi\lambda\left(1 + \frac{(x-\mu)^2}{\lambda^2}\right)}$$

$\mu \in \mathbb{R}$

$\lambda \in (0,\infty)$

$$F(x) = \frac{1}{\pi}\arctan\left(\frac{x-\mu}{\lambda}\right) + 1/2$$

- Laplace distr

$$Laplace_\lambda(B) = 1/2\,\lambda \int_B e^{-\lambda|x|}\, dx$$

$\lambda \in (0,\infty)$

$$F(x) = \begin{cases} 1/2 \exp(\lambda x) & : x < 0 \\ 1 - 1/2 \exp(\lambda x) & : x > 0 \end{cases}$$

- Normal distr

$$\mathcal{N}_{0,\mathbb{1}_{\mathbb{R}^d}}(B) = \frac{1}{(2\pi)^{d/2}} \int_B \exp\left(-1/2 \|x\|^2_{\mathbb{R}^d}\right) dx$$

$d \in \mathbb{N}$

$v \in \mathbb{R}^d$

$$\mathcal{N}_{v,Q}(B) = \mathcal{N}_{0,\mathbb{1}_{\mathbb{R}^d}}\left(\sqrt{Q}\, id_{\mathbb{R}^d} + v \in B\right)$$

$Q \in \mathbb{R}^{d\times d}$, nonnegative, symmetric

$$= \mathcal{N}_{0,\mathbb{1}_{\mathbb{R}^d}}\left(\{x \in \mathbb{R}^d : \sqrt{Q}\,x + v \in B\}\right)$$

$$A\,X(\omega) + b \sim \mathcal{N}_{Av+b,\,AQA^T} \quad \text{with } X(\omega) \sim \mathcal{N}_{v,Q}$$

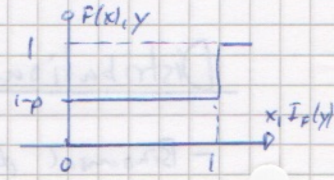
(generic) $\Phi(x) = 1/2 + 1/2\, erf\left(\frac{x-\mu}{\sqrt{2}\cdot\sigma}\right)$ $f(x) = \frac{1}{\sqrt{2\pi}\cdot\sigma} \cdot \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$

$\Phi(x) = 1/2 + 1/2\, erf\left(\frac{x}{\sqrt{2}}\right)$

- Bernoulli distr

$k \in \{0,1\}$ $\quad f(k) = \begin{cases} 1-p & : k=0 \\ p & : k=1 \end{cases}$ $\quad F(k) = \begin{cases} 0 & k<0 \\ 1-p & 0 \le k < 1 \\ 1 & 1 \le k \end{cases}$ $\quad I_F(y) = \begin{cases} 0 & : 0<y\le 1-p \\ 1 & : 1-p<y<1 \end{cases}$

$p_0 = 1-p, \; p_1 = p$



- Binomial distr

$k \in \{0,1,\dots,n\}$ $\quad f(k) = \binom{n}{k} p^k (1-p)^{n-k}$ $\quad F(k) = \sum_{i=0}^{\lfloor k \rfloor} \binom{n}{i} p^i (1-p)^{n-i}$

$= (n-k)\cdot\binom{n}{k} \int_0^{1-p} t^{n-k-1} (1-t)^k \, dt$

$p_{k+1} = \frac{p}{(1-p)} \cdot \frac{n-k}{k+1} \cdot p_k$

$p_0 = (1-p)^n$

- Geometric distr

$k = \{1,2,\dots\}$ # of trials for 1st success $\quad f(k) = (1-p)^{k-1} \cdot p$ $\quad F(k) = 1-(1-p)^k$ $\quad I_F(y) = \left\lceil \frac{\log(1-y)}{\log(1-p)} \right\rceil$ $\quad p_{k+1} = (1-p)\cdot p_k$

$k = \{0,1,\dots\}$ # of failures before 1st success $\quad f(k) = (1-p)^k \cdot p$ $\quad F(k) \;\; 1-(1-p)^{k+1}$ $\quad I_F(y) = \left\lceil \frac{\log(1-y)}{\log(1-p)} \right\rceil - 1$ $\quad p_0 = p$

- Poisson distr

$k \in \{0,1,\dots\}$ $\quad f(k) = \frac{\lambda^k}{k!} \cdot e^{-\lambda}$ $\quad F(k) = e^{-\lambda} \cdot \sum_{i=0}^{\lfloor k \rfloor} \frac{\lambda^i}{i!}$ $\quad p_0 = e^{-\lambda}$

$p_{n+1} = \frac{\lambda}{n+1} \cdot p_n$

- Uniform distr

$x \in [a,b]$ $\quad f(x) = \frac{1}{b-a}$ $\quad F(x) = \frac{x-a}{b-a}$ $\quad I_F(y) = a + y\cdot(b-a)$

- Exponential distr

$x \in [0,\infty)$ $\quad \lambda \in \mathbb{R}^+ \setminus \{0\}$ $\quad f(x) = \lambda \cdot e^{-\lambda x}$ $\quad F(x) = 1 - e^{-\lambda x}$ $\quad I_F(y) = \frac{-\log(1-y)}{\lambda}$

- Laplace distr

$\mu \in \mathbb{R}, \lambda \in \mathbb{R}^+ \setminus \{0\}, x \in \mathbb{R}$ $\quad f(x) = \tfrac{1}{2} \lambda e^{-\lambda |x-\mu|}$ $\quad F(x) = \begin{cases} \tfrac{1}{2} e^{\lambda(x-\mu)} & : x<\mu \\ 1-\tfrac{1}{2} e^{-\lambda(x-\mu)} & : x \ge \mu \end{cases}$ $\quad I_F(x) = \begin{cases} \mu + \frac{1}{\lambda}\log 2y & : 0<y<\tfrac{1}{2} \\ \mu - \frac{1}{\lambda}\log(2-2y) & : \tfrac{1}{2}<y<1 \end{cases}$

- Cauchy distr

$x \in \mathbb{R}, \mu \in \mathbb{R}, \lambda \in \mathbb{R} \setminus \{0\}$ $\quad f(x) = \frac{1}{\pi\lambda\cdot\left(1+\left(\frac{x-\mu}{\lambda}\right)^2\right)}$ $\quad F(x) = \frac{1}{\pi} \arctan\left(\frac{x-\mu}{\lambda}\right) + \tfrac{1}{2}$ $\quad I_F(y) = \lambda \cdot \tan(\pi\cdot(y-\tfrac{1}{2})) + \mu$

Generation of Random Numbers

- Inversion method

◦ generalized inverse distr func associated to a distr func

$$I_F(y) = \inf\{x \in \mathbb{R} : F(x) \ge y\} = \inf(F^{-1}([y,1]))$$

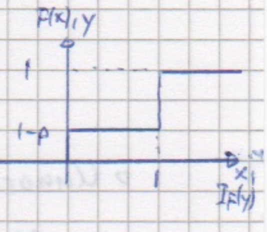
↳ properties: non-decreasing, i.e. $y_1 \le y_2 : I_F(y_1) \le I_F(y_2) \quad \forall\, y_1, y_2 \in (0,1)$

switching formula: $I_F(y) \le x \Leftrightarrow y \le F(x)$

$I_F(y) > x \Leftrightarrow y > F(x)$

◦ method: $\mathbb{P}[I_F(U) \le x] = F(x)$ with $U \sim \mathcal{U}_{(0,1)}$

▷ Bernoulli distr



$$F(x) = \begin{cases} 0 & : x < 0 \\ 1-p & : 0 \le x < 1 \\ 1 & : 1 \le x \end{cases} \qquad I_F = \begin{cases} 0 & : 0 < y \le 1-p \\ 1 & : 1-p < y < 1 \end{cases}$$

algorithm:
Output: Realization x of $X \sim Ber_p$
Generate Realization u of $U \sim \mathcal{U}_{(0,1)}$
if $u \le 1-p$ then $x=0$
else $x=1$
end if

or:
same output
Gen Real u of $U \sim \mathcal{U}_{(0,1)}$
if $u < p$ then $x=1$
else $x=0$
end if

in MATLAB: $\underbrace{\text{rand}(1,N) < p}_{\text{Real } x \text{ of } X \sim Ber_p}$ and $\underbrace{\text{sum}(\text{rand}(1,N) < p)/N}_{\text{Real } x \text{ of } X \sim b_{N,p}}$

▷ Discrete distr

$$I_F(y) = \min\{n \in \mathbb{N}_0 : F(n) \ge y\} = \min\{n \in \mathbb{N}_0 : \sum_{k=0}^{n} p_k \ge y\}$$

for any discrete variable $X(P)_{\mathcal{B}(\mathbb{R})} = \sum_{n=0}^{\infty} p_n \, \delta_n^{\mathbb{R}}\big|_{\mathcal{B}(\mathbb{R})}$

with $F(x) = \sum_{n=0}^{\lfloor x \rfloor} p_n$

▷ Poisson distr $\quad p_{n+1} = \frac{\lambda^{n+1}}{(n+1)!\, e^{\lambda}} = \frac{\lambda}{n+1} \cdot p_n$

algorithm:
Output: Real x of $X \sim Poi_\lambda$
Gen real u of $U \sim \mathcal{U}_{(0,1)}$
$n=0$
$p = e^{-\lambda}$
$F = p$
while $u > F$ do
  $p = p \cdot \lambda/(n+1)$
  $F = F + p$
  $n = n+1$
end while
$x = n$

▷ Binomial distr $\quad p_{k+1} = \binom{n}{k+1} p^{k+1} \cdot (1-p)^{n-(k+1)} = \frac{p}{1-p} \cdot \frac{n-k}{k+1} \cdot p_k$

algorithm:
Output: Real x of $X \sim b_{n,p}$
Gen real u of $U \sim \mathcal{U}_{(0,1)}$
$k=0$; $r = p/(1-p)$; $q = (1-p)^n$; $F = q$

```
while u>F do
    q = r·q·(n-k)/(k+1)
    F = F+q
    k = k+1
end while
x = k
```

- Geometric distr $p_n = p(1-p)^n$ $\quad F(n) = 1-(1-p)^{n+1}$

$$I_F = \left\lceil \frac{\log(1-u)}{\log(1-p)} \right\rceil_1 - 1 \qquad X = \left\lfloor \frac{\log u}{\log(1-p)} \right\rfloor_1$$

- Acceptance-Rejection method
  - Uniform distr

```
Output: Real x of X ~ U_A        => for A ⊆ B
Gen real y of Y ~ U_B
if y ∈ A then
    x = y   (Accept)
else
    Restart algorithm (Reject)
end if
```

  - Unnormalized density functions wrt. the Lebesgue-Borel measure

$$\mu(A) = \frac{\int_A f(x)\, \lambda_{\mathbb{R}^d}(dx)}{\int_{\mathbb{R}^d} f(x)\, \lambda_{\mathbb{R}^d}(dx)}$$

  - Acceptance-Rejection method : $(y, g(y)\cdot u) \in \text{subgraph}(f) \iff g(y)\cdot u \leq f(y)$

algorithm:

```
Output: Real x of X ~ X(P)_B(R^d)   (with unnormalized density f)
Gen real y of Y ~ Y(P)_B(R^d)       (  "        "         "   g)
Gen real u of U ~ U_(0,1)
if g(y)·u ≤ f(y)
    x = y   (Accept)
else
    restart algorithm (Reject)
end if
```

- Methods for the normal distr
  - central limit theorem (scalar case)

for Y_n an i.i.d. RV: $\lim_{n\to\infty} \mathbb{P}\left(\frac{Y_1+\dots+Y_n - n\cdot\mathbb{E}_{\mathbb{P}}[Y_1]}{\sqrt{n\cdot \mathrm{Var}_{\mathbb{P}}(Y_1)}} \leq x\right) = \mathcal{N}_{0,\mathbb{1}_{\mathbb{R}}}((-\infty, x])$

for $U_n \sim \mathcal{U}_{(0,1)}$, $\mathbb{E}[U_n] = 1/2$, $\mathrm{Var}_{\mathbb{P}}[U_n] = 1/12$, $S_n = \dots = \frac{U_1+\dots+U_n - \frac{1}{2}n}{\sqrt{n/12}}$

then $S_n \sim \mathcal{N}_{0,\mathbb{1}_{\mathbb{R}}}$ (for $n\in\mathbb{N}$ large)

algorithm:

```
Output: Real x of X ~ S_12 (P)_B(R) ≈ N_0,1
s = 0
for n = 1 -> 12 do
    Gen real u of U ~ U_(0,1)
    s = s+u
end for
x = s-6
```

in MATLAB:
S_{12} = sum(rand(1,12)) - 6

  - Box-Muller method

with $U_1, U_2 \sim \mathcal{U}_{(0,1)}$ $\Rightarrow$ $X_1 = \sqrt{-2\log(U_1)}\cdot\cos(2\pi U_2)$, $X_2 = \sqrt{-2\log(U_1)}\cdot\sin(2\pi U_2)$ $\Rightarrow X=(X_1,X_2)\sim\mathcal{N}_{0,\mathbb{1}_{\mathbb{R}^2}}$

  - Marsaglia Polar method

with $U:\Omega\to\mathbb{R}^2 \sim \mathcal{U}_{(x,y)\in\mathbb{R}^2:\, x^2+y^2\in(0,1)}$

$$X = \begin{pmatrix}U_1\\U_2\end{pmatrix}\cdot\frac{\sqrt{-2\log(\|U\|^2_{\mathbb{R}^2})}}{\|U\|_{\mathbb{R}^2}} \sim \mathcal{N}_{0,\mathbb{1}_{\mathbb{R}^2}}$$

```
Output: Real (x1,x2) of (X1,X2) ~ N_0,1
repeat
    Gen real (v1,v2) of (V1,V2) ~ U_(0,1)^2
    q = (2v1-1)^2 + (2v2-1)^2
until q ∈ (0,1)
w = sqrt(-2 log(q)/q)
x1 = (2v1-1)·w
x2 = (2v2-1)·w
```

- Normally distr RVs with general mean & covariance matrix
  - mean $v \in \mathbb{R}^d$, covariance $Q \in \mathbb{R}^{d \times d}$ & Q nonnegative, symmetric & $Q = A \cdot A^T$
  - with $X \sim \mathcal{N}_{0, \mathbb{1}_{\mathbb{R}^d}} \Rightarrow A \cdot X(\omega) + v \sim \mathcal{N}_{v, A \cdot A^T}$
  - MATLAB: chol(Q)' * randn(d,1) + v $\sim \mathcal{N}_{v,Q}$

Monte Carlo Integration methods

- Quadrature formula

$$Q[f] = \sum_{i \in I} w_i \cdot f(x_i) \qquad \text{quadrature nodes } (x_i)_{i \in I} \subseteq A \in \mathcal{B}(\mathbb{R}^d)$$
$$\text{" weights } (w_i)_{i \in I} \subseteq \mathbb{R}$$

- Rectangle method (d-dimensional left rectangle method)

$$R^n_{[a,b]^d} : \mathcal{L}^1(\mathcal{B}_{[a,b]^d}; |\cdot|_{\mathbb{R}}) \to \mathbb{R}, \quad n \in \mathbb{N}$$

$$R^n_{[a,b]^d}[f] = \frac{(b-a)^d}{n^d} \cdot \sum_{i_1,\dots,i_d \in \{0,1,\dots,n-1\}} f\left(a + \tfrac{i_1}{n}(b-a), \dots, a + \tfrac{i_d}{n}(b-a)\right) \approx \int_{a^d}^{b^d} f(x)\,dx$$

  - equivalent to the Quadrature formula with:
    quadrature nodes $\left(a + \frac{i_1}{n}(b-a), \dots, a + \frac{i_d}{n}(b-a)\right)$
    quadrature weights $\frac{(b-a)^d}{n^d}$
  - 1-dimensional formula: $R^n_{[a,b]} f = \frac{(b-a)}{n} \cdot \sum_{i=0}^{n-1} f\left(a + \frac{i}{n}(b-a)\right)$
  - error estimate

$$\left| R^n_{[a,b]^d}[f] - \int_{[a,b]^d} f(x)\,dx \right| \leq (b-a)^d \cdot w_f\left(\frac{(b-a)\cdot\sqrt{d}}{n}\right) \leq \frac{(b-a)^{d+\alpha} \cdot d^{\alpha/2} \cdot \|f\|_{C^{0,\alpha}([a,b]^d,\mathbb{R})}}{n^\alpha}$$

    with $d, n \in \mathbb{N}$, $\alpha \in (0,1]$, $a, b \in \mathbb{R}$ with $a < b$, $f \in \mathcal{L}^1(\mathcal{B}_{[a,b]^d}; |\cdot|_{\mathbb{R}})$
    i.e. if f is Hölder-continuous, then $\cdot \to 0$ & $R^n_{[a,b]^n}$ converges
    and: $R^n_{[a,b]^n}$ converges also if f is continuous

- 1-dimensional Trapezoidal method

$$T^n_{[a,b]} : \mathcal{L}^1(\mathcal{B}_{[a,b]}; |\cdot|_{\mathbb{R}}) \to \mathbb{R} \quad \text{and} \quad f \in \mathcal{L}^1(\mathcal{B}_{[a,b]}; |\cdot|_{\mathbb{R}})$$

$$T^n_{[a,b]} = \frac{b-a}{n} \cdot \left(\frac{f(a) + f(b)}{2} + \sum_{i=1}^{n-1} f\left(a + \tfrac{i}{n}(b-a)\right) \right)$$
$$= \frac{b-a}{n} \cdot \sum_{i=0}^{n-1} \tfrac{1}{2} \cdot \left(f\left(a + \tfrac{i}{n}(b-a)\right) + f\left(a + \tfrac{i+1}{n}(b-a)\right) \right)$$

  - error estimates $\alpha \in [0,1]$

$$\forall f \in \mathcal{L}^1(\mathcal{B}_{[a,b]}; |\cdot|_{\mathbb{R}}) : \left| T^n_{[a,b]}[f] - \int_a^b f(x)\,dx \right| \leq (b-a) \cdot w_f\left(\frac{b-a}{2n}\right) \leq \frac{(b-a)^{1+\alpha} \|f\|_{C^{0,\alpha}([a,b],\mathbb{R})}}{(2n)^\alpha}$$

$$\forall f \in C^1([a,b],\mathbb{R}) : \quad \text{"} \quad \leq \frac{(b-a)^2}{n} \cdot w_{f'}\left(\frac{b-a}{2n}\right) \leq \frac{(b-a)^{2+\alpha} \cdot \|f'\|_{C^{0,\alpha}([a,b],\mathbb{R})}}{2^\alpha \cdot n^{1+\alpha}}$$

  - Curse of dimensionality:
    Quadrature formulas may have a very poor convergence speed if:
    (i) the dimension $d \in \mathbb{N}$ is large
    (ii) the integrand function $f : [a,b]^d \to \mathbb{R}$ has low regularity properties

- Monte Carlo approximation:

$$\mathbb{E}_{\mathbb{P}}[X_1] \approx \frac{X_1 + \dots + X_N}{N} \qquad \text{with } X_n \in \mathcal{L}^1(\mathbb{P}; |\cdot|_{\mathbb{R}}) \text{ i.i.d. RVs}$$
$$= \int_A f(x)\,\mu(dx)$$

- Bias of an estimator
  - ▷ $\mathbb{P}$-bias of X wrt. c: $\quad \mathrm{Bias}_{\mathbb{P},c}(X) = \mathbb{E}_{\mathbb{P}}[X] - c$
  - ◦ $\mathbb{P}$-unbiased wrt. c: $\quad \mathrm{Bias}_{\mathbb{P},c}(X) = 0$
  - ◦ $\mathbb{P}$-biased wrt. c: $\quad \mathrm{Bias}_{\mathbb{P},c}(X) \neq 0$
  - ▷ Unbiasedness of Monte Carlo approximations:
    since $\mathbb{E}_{\mathbb{P}}[\frac{1}{N}\cdot(X_1 + \dots + X_N)] - \mathbb{E}_{\mathbb{P}}[X_1] = 0$
    $\Rightarrow \frac{1}{N}(X_1 + \dots + X_N)$ is $\mathbb{P}$-unbiased wrt. $\mathbb{E}_{\mathbb{P}}[X_1]$

- Consistency of the Monte Carlo method
  - ▷ almost sure convergence implies convergence in probability
    - the sequence $(X_n)_{n\in\mathbb{N}}$ converges $\mathbb{P}$-a.s. to X, i.e.
      $$\mathbb{P}\left[\limsup_{n\to\infty} |X_n - X|_{\mathbb{R}} = 0\right] = 1$$
    - this implies that the sequence $(X_n)_{n\in\mathbb{N}}$ also converges in probability to X, i.e.
      $$\limsup_{n\to\infty} \mathbb{P}\left[|X_n - X|_{\mathbb{R}} \geq \varepsilon\right] = 0 \quad \forall \varepsilon \in (0,\infty)$$
  - ▷ $(X_n)_{n\in\mathbb{N}}$ is $\mathbb{P}$-consistent for c if $(X_n)_{n\in\mathbb{N}}$ converges in probability to c, i.e.
    $$\limsup_{n\to\infty} \mathbb{P}\left[|X_n - X|_{\mathbb{R}} \geq \varepsilon\right] = 0 \quad \forall \varepsilon \in (0,\infty)$$
  - ◦ $(X_n)_{n\in\mathbb{N}}$ is strongly $\mathbb{P}$-consistent for c if $(X_n)_{n\in\mathbb{N}}$ converges $\mathbb{P}$-a.s. to c, i.e.
    $$\mathbb{P}\left[\limsup_{n\to\infty} |X_n - c|_{\mathbb{R}} = 0\right] = 1$$
  - ▷ Strong law of large numbers:
    Monte Carlo approximations $\frac{1}{N}(X_1 + \dots + X_N)$ are strongly $\mathbb{P}$-consistent for $\mathbb{E}_{\mathbb{P}}[X_1]$
    since the sequence $\frac{1}{N}(X_1 + \dots + X_N)$ converges $\mathbb{P}$-a.s. to $\mathbb{E}_{\mathbb{P}}[X_1]$, i.e.
    $$\mathbb{P}\left[\limsup_{N\to\infty} \left|\frac{1}{N}\cdot(X_1 + \dots + X_N) - \mathbb{E}_{\mathbb{P}}[X_1]\right|_{\mathbb{R}} = 0\right] = 1$$

- Root mean square error of the Monte Carlo method
  for $X_n \in \mathcal{L}^2(\mathbb{P}, |\cdot|_{\mathbb{R}})$: $\quad \left\| \mathbb{E}_{\mathbb{P}}[X_1] - \frac{1}{N}(X_1 + \dots + X_N)\right\|_{\mathcal{L}^2(\mathbb{P},|\cdot|_{\mathbb{R}})} = \frac{\sqrt{\mathrm{Var}_{\mathbb{P}}(X_1)}}{\sqrt{N}}$

- Markov inequality
  $(\Omega, \mathcal{F}, \mu)$ measure space
  $X: \Omega \to [0,\infty)$ $\mathcal{F}/\mathcal{B}([0,\infty))$-meas. $\Rightarrow \forall \varepsilon \in (0,\infty): \mu(X \geq \varepsilon) \leq \frac{\int_\Omega X \, d\mu}{\varepsilon}$
  ↳ it follows that L^p convergence (i.e. $\limsup_{n\to\infty} \|X_n\|_{\mathcal{L}^p(\mathbb{P},|\cdot|_{\mathbb{R}})} = 0$)
  implies convergence in probability (i.e. $\limsup_{n\to\infty} \mathbb{P}[|X_n|_{\mathbb{R}} \geq \varepsilon] = 0$)

- Chebyshev inequality
  $$\mathbb{P}\left[|X - \mathbb{E}_{\mathbb{P}}[X]|_{\mathbb{R}} \geq \varepsilon\right] \leq \frac{\mathbb{E}_{\mathbb{P}}[|X - \mathbb{E}_{\mathbb{P}}[X]|^q]}{\varepsilon^q}$$
  for: $\varepsilon, q \in (0,\infty)$, $\mathbb{E}_{\mathbb{P}}[|X|_{\mathbb{R}}] < \infty$
  $X: \Omega \to \mathbb{R}$ $\mathcal{F}/\mathcal{B}(\mathbb{R})$-meas.

  and $\quad \mathbb{P}\left[|X - \mathbb{E}_{\mathbb{P}}[X]|_{\mathbb{R}} \geq \varepsilon\right] \leq \frac{\mathrm{Var}_{\mathbb{P}}[X]}{\varepsilon^2}$

- Approximating the variance of a RV
  - exact variance: $\mathrm{Var}_{\mathbb{P}}[X_1] = \mathbb{E}_{\mathbb{P}}[(X_1 - \mathbb{E}_{\mathbb{P}}[X_1])^2]$
  - if $\mathbb{E}_{\mathbb{P}}[X_1]$ is known:
    $$\bar{X}_n = (X_n - \mathbb{E}_{\mathbb{P}}[X_1])^2 \Rightarrow \frac{\bar{X}_1 + \dots + \bar{X}_N}{N} \approx \frac{1}{N}\sum_{n=1}^{N}(X_n - \mathbb{E}_{\mathbb{P}}[X_1])^2$$
    ↳ which is $\mathbb{P}$-unbiased wrt. $\mathrm{Var}_{\mathbb{P}}[X_1]$
  - if $\mathbb{E}_{\mathbb{P}}[X_1]$ is unknown (as usual):
    $$\frac{1}{N}\sum_{n=1}^{N}\left(X_n - \frac{X_1 + \dots + X_N}{N}\right)^2 \quad \text{is } \mathbb{P}\text{-biased wrt. } \mathrm{Var}_{\mathbb{P}}[X_1]$$
    $$\frac{1}{N-1}\sum_{n=1}^{N}\left(X_n - \frac{X_1 + \dots + X_N}{N}\right)^2 \quad \text{is } \mathbb{P}\text{-unbiased wrt. } \mathrm{Var}_{\mathbb{P}}[X_1]$$
    since:
    $$\mathrm{Var}_{\mathbb{P}}[X_1] = \mathbb{E}_{\mathbb{P}}\left[\frac{1}{N-1}\sum_{n=1}^{N}\left(X_n - \frac{X_1 + \dots + X_N}{N}\right)^2\right] = \mathbb{E}_{\mathbb{P}}\left[\frac{1}{N}\sum_{n=1}^{N}(X_n - \mathbb{E}_{\mathbb{P}}[X_1])^2\right]$$

Continuity Properties

Continuously differentiable $\subseteq$ Lipschitz cont. $\subseteq$ α-Hölder cont. $\subseteq$ uniformly cont $\subseteq$ cont.

- Hölder continuity
  - (E, d_E), (F, d_F) metric spaces, $\alpha \in [0,1]$
  - we denote by
    $$|\cdot|_{C^\alpha(E,F)} : \mathbb{M}(E,F) \to [0,\infty]$$
    the mapping with the property that $\forall f \in \mathbb{M}(E,F)$ it holds that
    $$|f|_{C^\alpha(E,F)} = \sup_{\substack{x,y \in E \\ x \neq y}} \left[\frac{d_F(f(x), f(y))}{|d_E(x,y)|^\alpha} \right] \in [0,\infty]$$
  - and we denote by $C^\alpha(E,F)$ the set given by
    $$C^\alpha(E,F) = \{ f \in \mathbb{M}(E,F) : |f|_{C^\alpha(E,F)} < \infty \}$$
  - If $f \in C^\alpha(E,F)$, then we say that f is α-Hölder continuous.
  - explanations: · $\mathbb{M}(A,B)$ denotes the set of all mapping from A to B, where A, B are sets
    · the metric space (E, d_E) is given by
      a set E
      & a mapping (distance function) $d_E : E \times E \to [0,\infty)$
  - sample paths of BM are a.s. everywhere locally α-Hölder cont. with $\alpha < 1/2$
  - if $\alpha = 1 \Rightarrow f$ is Lipschitz cont.
  - if $\alpha = 0 \Rightarrow f$ is simply bounded

- Lipschitz continuity
  - (E, d_E), (F, d_F) metric spaces, $f : E \to F$
  - f is called Lipschitz continuous if $\exists K \geq 0$ s.t.
    $$d_F(f(x), f(y)) \leq K \cdot d_E(x,y) \quad \forall x,y \in E,\ x \neq y$$
    or: $$\frac{d_F(f(x), f(y))}{d_E(x,y)} \leq K$$
  - Lipschitz continuous functions:
    $f(x) = \sqrt{x^2+5}$ $(K=1)$ $\quad f(x) = |x|$ $(K=1)$
    $f(x) = \sin(x)$
  - Not (globally) Lipschitz cont. functions:
    $f(x) = \sqrt{x}$ on $[0,1]$ (since it becomes infinitely steep at $x=0$)
    (but it is Hölder cont. of class $C^{0,\alpha}$ for $\alpha \leq 1/2$)
    $f(x) = x^{3/2} \cdot \sin(1/x)$
    $f(x) = e^x$ (since it becomes arbitrarily steep as $x \to \infty$; but it is locally Lipschitz cont.)
    $f(x) = x^2$ (only locally Lipschitz cont.)

Stochastic processes & Itô stochastic calculus & SDEs

- Stochastic integral
  - $I^W_{a,b}(X) = \int_a^b X_s \, dW_s$
  - Properties:

$$\mathbb{E}_{\mathbb{P}}\left[\int_a^b Y_s \, dW_s\right] = 0$$

$$\mathbb{E}_{\mathbb{P}}\left[\left\| \int_a^b Y_s \, dW_s \right\|^2_{\mathbb{R}^d}\right] = \int_a^b \mathbb{E}_{\mathbb{P}}\left[\|Y_s\|^2_{HS(\mathbb{R}^m, \mathbb{R}^d)}\right] ds \qquad \text{(Itô's isometry)}$$

$$\left\| \int_a^b Y_s \, dW_s \right\|_{L^2(\mathbb{P}; \|\cdot\|_{\mathbb{R}^d})} = \left(\int_a^b \|Y_s\|^2_{L^2(\mathbb{P}; \|\cdot\|_{HS(\mathbb{R}^m, \mathbb{R}^d)})} ds \right)^{1/2}$$

...

- Itô process

drift Y, diffusion Z

$$X_t = X_0 + \int_0^t Y_s \, ds + \int_0^t Z_s \, dW_s$$

$$dX_t = Y_t \, dt + Z_t \, dW_t$$

- Itô's formula

$$f(X_t) = f(X_{t_0}) + \int_{t_0}^t f'(X_s) Y_s \, ds + \int_{t_0}^t f'(X_s) Z_s \, dW_s + \tfrac{1}{2} \sum_{i=1}^m \int_{t_0}^t f''(X_s)\left(Z_s e_i^{(m)}, Z_s e_i^{(m)}\right) ds$$

- Solution processes of SDEs

$X_t = \xi + \int_0^t \mu(X_s) \, ds + \int_0^t \sigma(X_s) \, dW_s$ is a solution process to the SDE: $dX_t = \mu(X_t) \, dt + \sigma(X_t) \, dW_t$, $t \in [0, T]$, $X_0 = \xi$

  - Indistinguishability: Local Lipschitz continuity of μ & σ is a sufficient but not a necessary condition to ensure that the solution processes of an SDE are unique up to indistinguishability.
  - Existence & Uniqueness:
    - Assume $\mu: \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma: \mathbb{R}^d \to \mathbb{R}^{d \times m}$ are globally Lipschitz cont., i.e. $\exists$ $C \in [0, \infty)$ s.t. $\forall$ $x, y \in \mathbb{R}^d$

$$\|\mu(x) - \mu(y)\|_{\mathbb{R}^d} + \|\sigma(x) - \sigma(y)\|_{\mathbb{R}^d} \leq C \cdot \|x - y\|_{\mathbb{R}^d}$$

    Then:

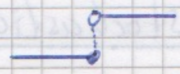
    - $\exists$ an up to indistinguishability unique solution process $X: [0,T] \times \Omega \to \mathbb{R}^d$ of the SDE

$$dX_t = \mu(X_t) \, dt + \sigma(X_t) \, dW_t, \quad t \in [0,T], \quad X_0 = \xi$$

    - it holds that

$$\sup_{t \in [0,T]} \|X_t\|_{L^p(\mathbb{P}; \|\cdot\|_{\mathbb{R}^d})} < \infty$$

    - it holds $\forall$ $\alpha \in (0, 1/2]$ that $X \in C^\alpha([0,T], L^p(\mathbb{P}; \|\cdot\|_{\mathbb{R}^d}))$

- Adaptivity with cont. sample paths implies predictability
  ↳ left-continuity is sufficient 

Numerical Schemes

- Approximative simulation of standard BM
  - scheme: $W_{t+1} = W_t + \sqrt{T/N} \cdot RV$ with $RV \sim \mathcal{N}_{0,1}$
  - linear interpolation: for $t \in \left[\frac{nT}{N}, \frac{(n+1)T}{N}\right]$

    $$\tilde{W}_t = \left(n+1-N\tfrac{t}{T}\right) W_{\frac{nT}{N}} + \left(N\tfrac{t}{T}-n\right) W_{\frac{(n+1)T}{N}}$$

- Euler-Maruyama scheme
  - SDE: $dX_t = \mu(X_t)\,dt + \sigma(X_t)\,dW_t \qquad t \in [0,T] \qquad X_0 = \xi$

    $X_t \approx X_{t_0} + \mu(X_{t_0})(t-t_0) + \sigma(X_{t_0})(W_t - W_{t_0})$

    ↳ only for $t-t_0$ "sufficiently small"
  - scheme: $Y_{n+1} = Y_n + \mu(Y_n)\cdot\frac{T}{N} + \sigma(Y_n)\cdot\left(W_{\frac{n+1}{N}T} - W_{\frac{n}{N}T}\right), \quad Y_0 = \xi$

    with: $T \in (0,\infty) \qquad \mu \in \mathcal{M}(\mathcal{B}(\mathbb{R}^d), \mathcal{B}(\mathbb{R}^d)) \qquad \xi \in \mathcal{L}^0(|\mathcal{F}_0; \|\cdot\|_{\mathbb{R}^d})$

    $d,m,N \in \mathbb{N} \qquad \sigma \in \mathcal{M}(\mathcal{B}(\mathbb{R}^d), \mathcal{B}(\mathbb{R}^{d\times m})) \qquad W: [0,T]\times\Omega \to \mathbb{R}^m$

    $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\in[0,T]})$ a stochastic basis

    $Y: \{0,1,\dots,N\}\times\Omega \to \mathbb{R}^d$
  - linear interpolation: for $t \in \left[\frac{n}{N}T, \frac{n+1}{N}T\right]$:

    $$Y_t = Y_{\frac{n}{N}T} + \left(\tfrac{t}{T}N - n\right)\cdot\left(\mu(Y_{\frac{n}{N}T})\cdot\tfrac{T}{N} + \sigma(Y_{\frac{n}{N}T})\cdot\left(W_{\frac{n+1}{N}T} - W_{\frac{n}{N}T}\right)\right)$$
  - convergence: · strong convergence · almost sure convergence

    · uniform strong convergence · numerically weak convergence
  - drift-implicit scheme:

    $$Y_{n+1} = Y_n + \mu(Y_{n+1})\tfrac{T}{N} + \sigma(Y_n)\left(W_{\frac{n+1}{N}T} - W_{\frac{n}{N}T}\right)$$
  - increment-tamed scheme:

    $$Y_{n+1} = Y_n + \frac{\mu(Y_n)\cdot\frac{T}{N} + \sigma(Y_n)\cdot\left(W_{\frac{n+1}{N}T} - W_{\frac{n}{N}T}\right)}{\max\left(1, \frac{T}{N}\cdot\left\|\mu(Y_n)\cdot\frac{T}{N} + \sigma(Y_n)\cdot\left(W_{\frac{n+1}{N}T} - W_{\frac{n}{N}T}\right)\right\|_{\mathbb{R}^d}\right)}$$

  } apply to SDEs with superlinearly growing coefficients $\mu(x), \sigma(x)$

- Milstein scheme
  - scheme: $Y_{n+1} = Y_n + \mu(Y_n)\frac{T}{N} + \sigma(Y_n)\left(W_{\frac{n+1}{N}T} - W_{\frac{n}{N}T}\right)$

    $$+ \sum_{i,j=1}^{m} \sigma_i'(Y_n)\cdot\sigma_j(Y_n)\cdot \underbrace{\int_{\frac{n}{N}T}^{\frac{n+1}{N}T}\int_{\frac{n}{N}T}^{s} dW_u^{(i)}\, dW_s^{(j)}}_{}$$

    $$= \tfrac{1}{2}\left(W_{\frac{n+1}{N}T} - W_{\frac{n}{N}T}\right)^2 - \tfrac{1}{2}\left(\tfrac{n+1}{N}T - \tfrac{n}{N}T\right)$$

    $$= \tfrac{1}{2}\Delta W^2(h) - \tfrac{1}{2}h$$
  - with commutative noise: condition: $\sigma_i'(x)\cdot\sigma_j(x) = \sigma_i(x)\cdot\sigma_j'(x) \quad \forall x \in \mathbb{R}^d \quad \forall i,j \in \{1,2,\dots,m\}$

    then: $Y_{n+1} = Y_n + \mu(Y_n)\frac{T}{N} + \sigma(Y_n)\cdot\left(W_{\frac{n+1}{N}T} - W_{\frac{n}{N}T}\right) - \frac{T}{2N}\sum_{i=1}^{m}\sigma_i'(Y_n)\cdot\sigma_i(Y_n)$

    $$\underbrace{+ \tfrac{1}{2}\sigma'(Y_n)\cdot\left(\sigma(Y_n)\cdot\left(W_{\frac{n+1}{N}T} - W_{\frac{n}{N}T}\right)\right)\cdot\left(W_{\frac{n+1}{N}T} - W_{\frac{n}{N}T}\right)}_{}$$

    $$= \tfrac{1}{2}\sum_{i,j=1}^{m}\sigma_i'(Y_n)\cdot\sigma_j(Y_n)\cdot\left(W^{(i)}_{\frac{n+1}{N}T} - W^{(i)}_{\frac{n}{N}T}\right)\cdot\left(W^{(j)}_{\frac{n+1}{N}T} - W^{(j)}_{\frac{n}{N}T}\right)$$

- Notions of convergence for stochastic processes
  - $(Y^N)_{N\in\mathbb{N}}$ converges at time T in the strong L^p-sense to Y^0 if:
    $\limsup_{N\to\infty} \mathbb{E}[\| Y_T^0 - Y_T^N \|_{\mathbb{R}^d}] = 0$
  - $(Y^N)_{N\in\mathbb{N}}$ converges at time T in the strong L^p-sense with order α to Y^0 if:
    $\| Y_T^0 - Y_T^N \|_{L^p(\mathbb{P};\|\cdot\|_{\mathbb{R}^d})} \le \frac{C}{N^\alpha} \quad C \in \mathbb{R}$
  - $(Y^N)_{N\in\mathbb{N}}$ converges at time T $\mathbb{P}$-a.s. to Y^0 if:
    $\mathbb{P}\left[\limsup_{N\to\infty} \| Y_T^0 - Y_T^N \|_{\mathbb{R}^d} = 0\right] = 1$
  - $(Y^N)_{N\in\mathbb{N}}$ converges at time T $\mathbb{P}$-a.s. with order α to Y^0 if:
    $\| Y_T^0 - Y_T^N \|_{\mathbb{R}^d} \le \frac{C}{N^\alpha}$ $\mathbb{P}$-a.s.
  - $(Y^N)_{N\in\mathbb{N}}$ converges at time T in probability to Y^0 if:
    $\limsup_{N\to\infty} \mathbb{P}[\| Y_T^0 - Y_T^N \|_{\mathbb{R}^d} \ge \epsilon] = 0 \quad \forall \epsilon \in (0,\infty)$
  - $(Y^N)_{N\in\mathbb{N}}$ converges at time T in the numerically weak sense to Y_0 if:
    $\limsup_{N\to\infty} | \mathbb{E}[\varphi(Y_T^0)] - \mathbb{E}[\varphi(Y_T^N)] |_{\mathbb{R}} = 0$
    - for every infinitely often differentiable function $\varphi: \mathbb{R}^d \to \mathbb{R}$ with at most polynomially growing derivatives
    - it furthermore holds that $\mathbb{E}[|\varphi(Y_T^0)|_{\mathbb{R}}] < \infty$
  - $(Y^N)_{N\in\mathbb{N}}$ converges at time T in the numerically weak sense with order α to Y^0 if:
    $| \mathbb{E}[\varphi(Y_T^0)] - \mathbb{E}[\varphi(Y_T^N)] |_{\mathbb{R}} \le \frac{C}{N^\alpha}, \quad C \in \mathbb{R}$
    - for every φ as above
    - it furthermore holds that $\mathbb{E}[|\varphi(Y_T^0)|_{\mathbb{R}} + |\varphi(Y_T^N)|_{\mathbb{R}}] < \infty$
- Convergence of the Euler-Maruyama scheme
  If the coefficients $\mu(x)$ & $\sigma(x)$ of the SDE are globally Lipschitz cont.,
  ↳ e.g. $\mu(x), \sigma(x)$ grow at most linearly
  then the Euler-Maruyama scheme features:
  - strong convergence (L^p) with order $\alpha \in (0, 1/2]$
  - uniform strong convergence
  - almost sure convergence with order $\alpha \in (0, 1/2)$
  - (imply) convergence in probability
  - numerically weak convergence with order $\alpha = 1$
- Divergence of the Euler-Maruyama scheme
  If the drift & diffusion coefficients $\mu(x)$ & $\sigma(x)$ are superlinearly growing,
  ↳ e.g. all polinomials with degree ≥ 2
  then the Euler-Maruyama scheme:
  - diverges in the strong L^p-sense
  - diverges in the numerically weak sense
- Strong convergence of the Milstein scheme
  If $\mu(x), \sigma(x)$ are twice continuously differentiable functions with globally bounded derivatives, then the linearly interpolated Milstein approximations $(Y^N)_{N\in\mathbb{N}}$ converge at time T in the strong L^p-sense with order 1 to X, i.e.
  $$\left(\mathbb{E}[\| X_T - Y_T^N \|_{\mathbb{R}^d}]\right)^{1/p} = \| X_T - Y_T^N \|_{L^p(\mathbb{P};\|\cdot\|_{\mathbb{R}^d})} \le \frac{C}{N}$$

Examples of SDEs

- Geometric BM

$$dX_t = \alpha \cdot X_t \cdot dt + \beta \cdot X_t \cdot dW_t \quad , t \in [0,T] \; , X_0 = \xi \quad , \alpha, \beta \in \mathbb{R}$$

$$\hookrightarrow X_t = \exp\left((\alpha - \tfrac{1}{2}\beta^2) \cdot t + \beta \cdot W_t \right) \cdot \xi \quad , \mathbb{E}[X_t] = \xi \cdot e^{\alpha \cdot t}$$

- Black-Scholes model

$$\begin{cases} dX_t^{(1)} = r \cdot X_t^{(1)} dt \\ dX_t^{(2)} = \alpha \cdot X_t^{(2)} dt + \beta \cdot X_t^{(2)} dW_t \end{cases} \quad , t \in [0,T] \; , X_0 = \xi \; , r, \alpha \in \mathbb{R}, \beta \in (0,\infty)$$

$x = (x_1, x_2) \in (0,\infty)^2$

$$\hookrightarrow \begin{cases} X_t^{(1)} = e^{r t} \cdot \xi^{(1)} \\ X_t^{(2)} = \xi^{(2)} \cdot \exp\left((\alpha - \tfrac{1}{2}\beta^2) \cdot t + \beta \cdot W_t \right) \end{cases}$$

$\triangleright$ derivative pricing:

$$D_T = f(X^{(2)}) = (X^{(2)} - K)^+$$

$\to D_0 = e^{-rT} \cdot \mathbb{E}[f(\tilde{X})]$ where $\tilde{X}$ a solution to $d\tilde{X}_t = r \cdot \tilde{X}_t \cdot dt + \beta \cdot \tilde{X}_t \cdot dW_t$

then: if $K > 0$:

$$D_0 = X_0^{(2)} \cdot \Phi\left(\frac{(r + \tfrac{1}{2}\beta^2) \cdot T + \log X_0^{(2)}/K}{\beta \sqrt{T}} \right) - K \cdot e^{-rT} \cdot \Phi\left(\frac{(r - \tfrac{1}{2}\beta^2) T + \log X_0^{(2)}/K}{\beta \sqrt{T}} \right)$$

if $K \leq 0$: $D_0 = X_0^{(2)} - K \cdot e^{-rT}$

- Stochastic Ginzburg-Landau equation

$$\mu(x) = \alpha \cdot x - \delta x^3 \qquad \sigma(x) = \beta x + \tilde{\beta}$$

$$dX_t = (\alpha \cdot X_t - \delta \cdot X_t^3) \cdot dt + (\beta \cdot X_t + \tilde{\beta}) \cdot dW_t \quad , t \in [0,T] \quad , X_0 = \xi$$

$\Rightarrow$ Although $\mu(x)$ is not globally Lipschitz cont., there exists an up to indistinguishability unique solution process

$\Rightarrow$ in addition, it holds that $\sup_{t \in [0,T]} \mathbb{E}[\,|X_t|_{\mathbb{R}}^p\,] < \infty \quad \forall p \in [0,\infty)$

- Stochastic Verhulst equation

$$\mu(x) = (\eta + \tfrac{1}{2}c^2) x - \lambda x^2 \qquad \sigma(x) = c \cdot x \quad ; c \in \mathbb{R} \; ; \eta, \lambda \in (0,\infty)$$

$$dX_t = \left((\eta + \tfrac{1}{2}c^2) \cdot X_t - \lambda X_t^2 \right) \cdot dt + c \cdot X_t \cdot dW_t \quad , t \in [0,T] \quad , X_0 = \xi$$

$\Rightarrow$ Although $\mu(x)$ is not globally Lipschitz cont., $\exists$ an up to indist. unique solution
&

$$\sup_{t \in [0,T]} \mathbb{E}[\,|X_t|_{\mathbb{R}}^p\,] < \infty$$

- Stochastic predator-prey model

$$\mu\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1(\alpha - \beta x_2) \\ x_2(\gamma x_1 - \delta) \end{pmatrix} \qquad \sigma\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} c_1 x_1 & 0 \\ 0 & c_2 x_2 \end{pmatrix} \quad ; \quad \begin{array}{l} c_1, c_2 \in \mathbb{R} \\ \alpha, \beta, \gamma, \delta \in (0,\infty) \end{array}$$

$$dX_t = \begin{pmatrix} X_t^{(1)} (\alpha - \beta \cdot X_t^{(2)}) \\ X_t^{(2)} (\gamma \cdot X_t^{(1)} - \delta) \end{pmatrix} dt + \begin{pmatrix} c_1 \cdot X_t^{(1)} & 0 \\ 0 & c_2 \cdot X_t^{(2)} \end{pmatrix} dW_t \quad , t \in [0,T] \; , X_0 = \xi$$

$\Rightarrow$ Althoug $\mu(x)$ not glob. Lip. cont., $\exists$ up to ind. unique solution

& $\sup_{t \in [0,T]} \mathbb{E}[\,|X_t|_{\mathbb{R}}^p\,] < \infty$