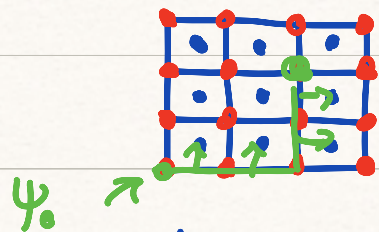
Lecture 15: Correlated Random Fields

Logistics: - HW5 is due Thursday

⇒ last chance on HW3 !

Last time: - Numerical Streamfunction

$$\psi = \psi_0 - \int q_y \, dx + \int q_x \, dy$$



- located in corners
- cumsum

- Choose ref. point & order of integration
  (ref. point affects the constant)

Today: - Correlated random fields

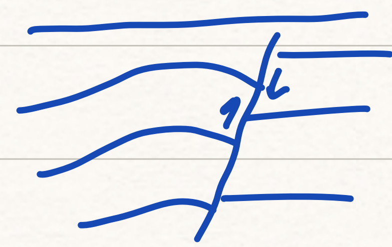
⇒ generate heterogeneous fields

⇒ multiple realizations → uncertainty

Generating Correlated Random Fields

subsurface is heterogeneous at all scales!

- large scale geological structure
- smaller scale random variation within each geological unit



Both are important, but the random component introduces uncertainty even if large-scale structure is known.

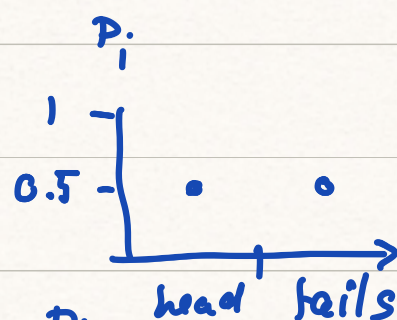
⇒ generate large numbers of distinct ~~ra~~ correlated random fields with the same statistics.

⇒ Geostatistics (PGE 337)

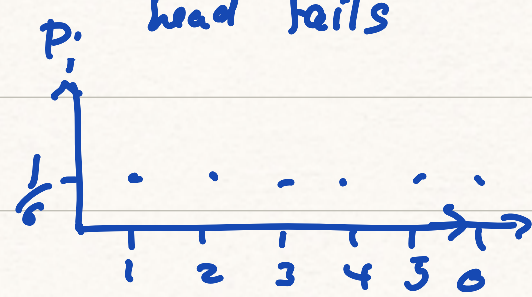
Some basic language:

Random variable X:

- discrete random var:
  - coin toss

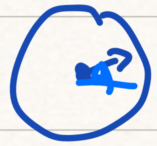


  - dice roll

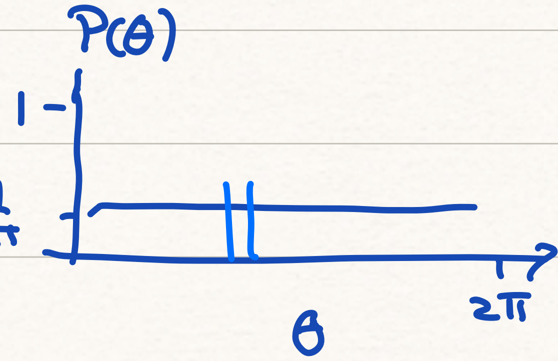


⇒ discrete events X_i discrete prob. P_i

• continuous random variables



"spinner" can land on any location



⇒ Continuous probability density function

Parameters describing aquifers are continuous.

Properties of random variables

1) Expected value (mean)

discrete: $E(X) = \sum_{i=1}^{N} X_i P_i = \mu$

X_i = ith outcome

P_i = probability of ith outcome

continuous: $E(X) = \int X \, P(X) \, dX = \mu$

2) Variance

squared deviation from mean

$$\mathrm{Var}(X) = E[(X-\mu)^2] = E[X^2] - E[X]^2 = \sigma^2$$

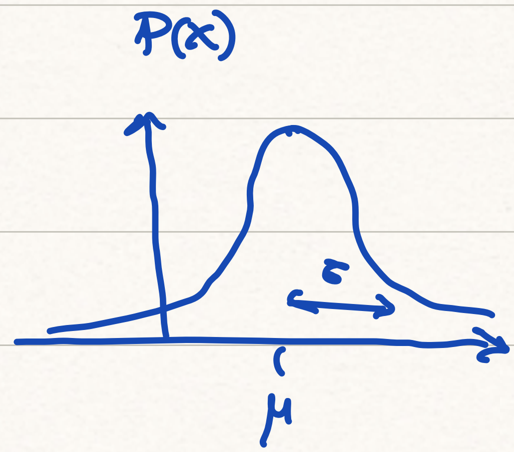
σ = standard deviation

Discrete: $\mathrm{Var}(X) = \sum_{i=1}^{n} (x_i - \mu)^2 p_i$

Continuum: $\mathrm{Var}(X) = \int (x-\mu)^2 P(x)\, dX$

Typical probability density functions

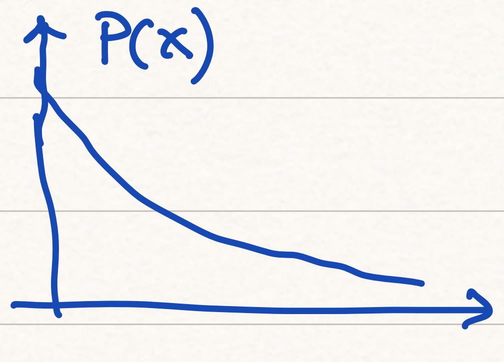
Normal distribution:



$$P(X) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

μ = mean $\quad$ σ = std. dev

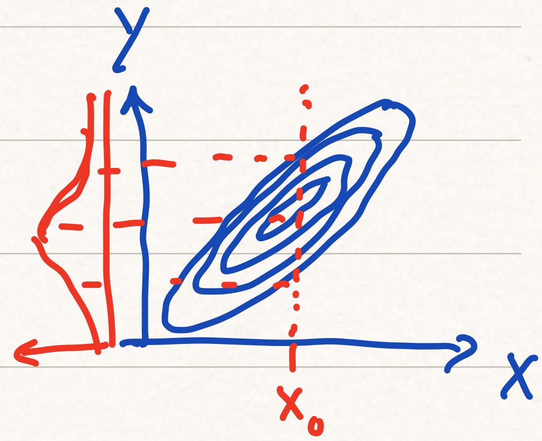
Exponential distribution



$$P(X) = \lambda e^{-\lambda X} \qquad X \geq 0$$

$$E(X) = \mu = \frac{1}{\lambda} \qquad \mathrm{Var}(X) = \frac{1}{\lambda^2}$$

Covariance & Correlation



If two random variables X and Y are not independent, i.e., they are jointly distributed (p_{ij}, $P(X,Y)$) we can compute their covariance

$$Cov(X,Y) = E[(X-\mu_X)(Y-\mu_Y)]$$

Discrete: $Cov(X,Y) = \sum_{i=1}^{N}\sum_{j=1}^{N}(x_i-\mu_X)(y_j-\mu_Y)\,p_{ij}$

Continuous: $Cov(X,Y) = \iint (X-\mu_X)(Y-\mu_Y)\,P(X,Y)\,dX\,dY$

$$Cov(X,X) = Var(X) = \sigma^2$$

Correlation: $\boxed{\rho_{X,Y} = \frac{Cov(X,Y)}{\sigma_X\,\sigma_Y}}$

Random Fields / Functions

So far no spatial extent!

Random field: $Z = \{ Z(\underline{x}) : x \in \mathbb{R}^d \}$

where $Z(\underline{x})$ is a scalar random variable at location $\underline{x}$.

If the field is stationary, i.e. statistics (E, Var) do not depend on $\underline{x}$ then we can define covariance

$$C(\underline{h}) = \mathrm{Cov}(Z(\underline{x}), Z(\underline{x} + \underline{h})) \qquad \underline{h} \in \mathbb{R}^d$$

where $\underline{h} = \underline{x} - \underline{x}$ is the lag vector ($|\underline{h}|$ is distance)

For an isotopic random field the covariance is only a function of distance $h = |\underline{h}|$

Correlation function: $\boxed{\rho(h) = \frac{C(h)}{\sigma^2}}$

Commonly used isotropic correlation functions:

1) Power exponential:

$$\rho(h) = \exp\left(-\left(\frac{h}{\kappa}\right)^{\nu}\right), \quad \kappa > 0, \; 0 \leq \nu \leq 2$$

$\nu = 2 \Rightarrow$ gaussian/normal

2) Rational quadratic (Cauchy)

$$\rho(h) = \frac{1}{\left(1 + \left(\frac{h}{\kappa}\right)^2\right)^{\nu}} \qquad \kappa > 0, \; \nu > 0$$

3) Matérn

$$\rho(h) = \frac{1}{\Gamma(\nu)\, 2^{\nu-1}} \left(\frac{2\sqrt{\nu}\, h}{\kappa}\right)^{\nu} K_{\nu}\left(\frac{2\sqrt{\nu}\, h}{\kappa}\right), \quad \begin{matrix}\kappa > 0 \\ \nu > 0\end{matrix}$$

K_ν mod. Bessel function

Γ Gamma function

Matern covariance fields are a parameterized family of fields that contain the most commonly used functions as limiting cases.

$\nu = \frac{1}{2}$ exponential correlation function

$\nu \to \infty$ ~~Ga~~ Normal correlation function

κ controlls how fast correlation decays.

Generating a random field

Form covariance matrix

$C_{ij} = \mathrm{Cov}(\underline{x}_i, \underline{x}_j)$ ⇒ given by correlation

Standard sampling of field with specified $\underline{\underline{C}}$ is as follows:

$$\boxed{\underline{m} = \underline{\underline{L}}\,\underline{s} + \mu}$$

μ = mean

$\underline{s}$ = white noise

$\underline{\underline{L}}$ = Cholesky decomp. of $\underline{\underline{C}}$ so that $\underline{\underline{C}} = \underline{\underline{L}}^T \underline{\underline{L}}$

Requires that we compute the Cholesky decomp.

⇒ both $\underline{\underline{C}}$ and $\underline{\underline{L}}$ are not sparse limits the size of fields that can be generated.

$\Rightarrow$ many ways of getting around this in Geostatistics

Relation between Matérn correlation functions and Stochastic PDE

The real benefit of Matern covariances is that they have been linked explicitly to solutions of SPDE. (Lindgren et. al. 2011)

$$\boxed{\left(-\nabla^2 + \kappa^2\right)^{\alpha/2} m = s}$$

where $s =$ white noise Gaussian random field with unit variance

$\alpha = \nu + \frac{d}{2}$

$d =$ dimension

$m =$ unknown parameter field

Our Matérn parameterization

choose $\alpha = 2$

$$\boxed{(-\gamma \nabla^2 + \delta)\, m = s}$$ mod. Helmholtz equ

two parameters γ & δ

We can show: $\rho \sim 2\sqrt{\frac{\gamma}{\delta}}$ $\sigma^2 \sim \frac{1}{\gamma\delta}$

Discretization:

$$\underbrace{(-\gamma\, \underline{\underline{D}} * \underline{\underline{G}} + \delta\, \underline{\underline{I}})}_{\underline{\underline{L}}}\, \underline{m} = \underline{s}$$

solve with all natural Neumann BC's