# Cost-effective Management of Remediation Projects

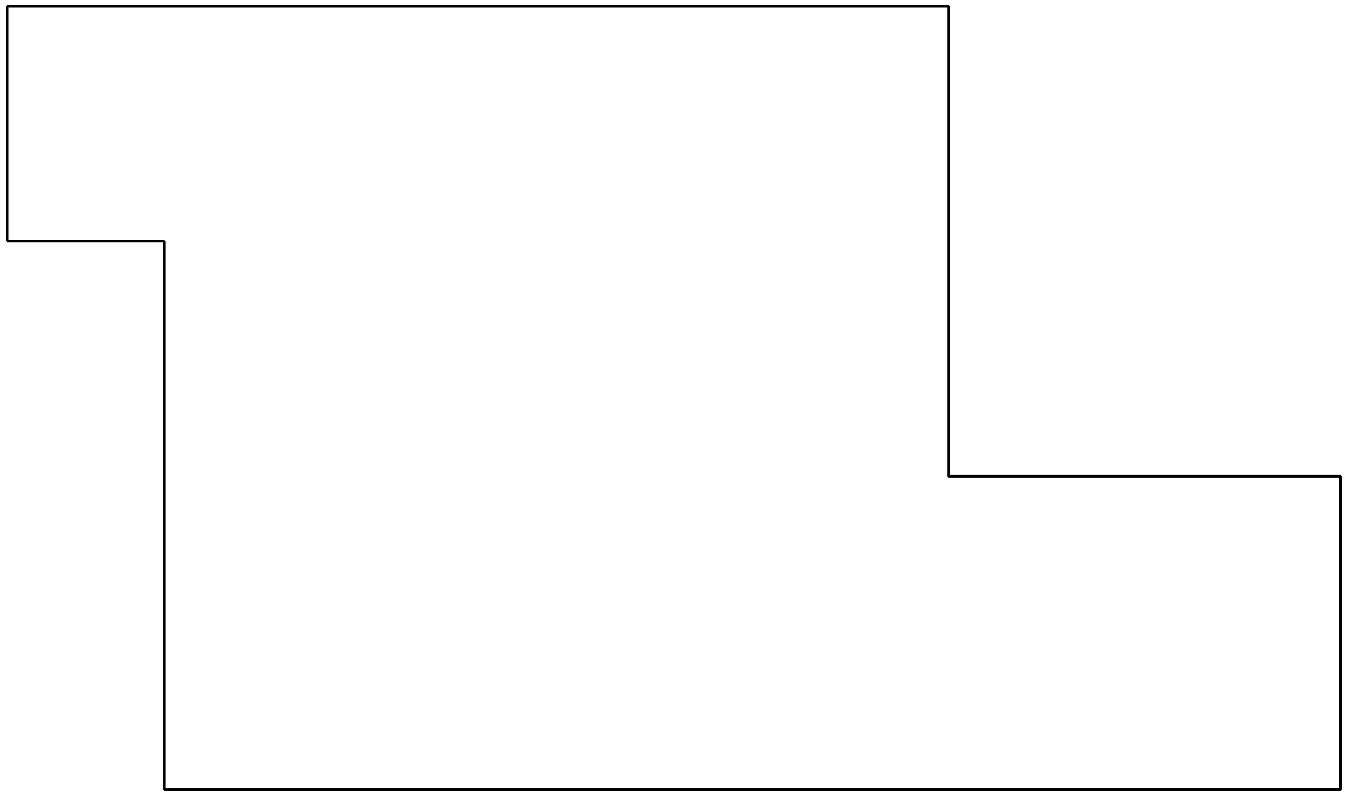
Tommy Norberg & Lars Rosén

Mathematical Statistics and Department of GeoEngineering

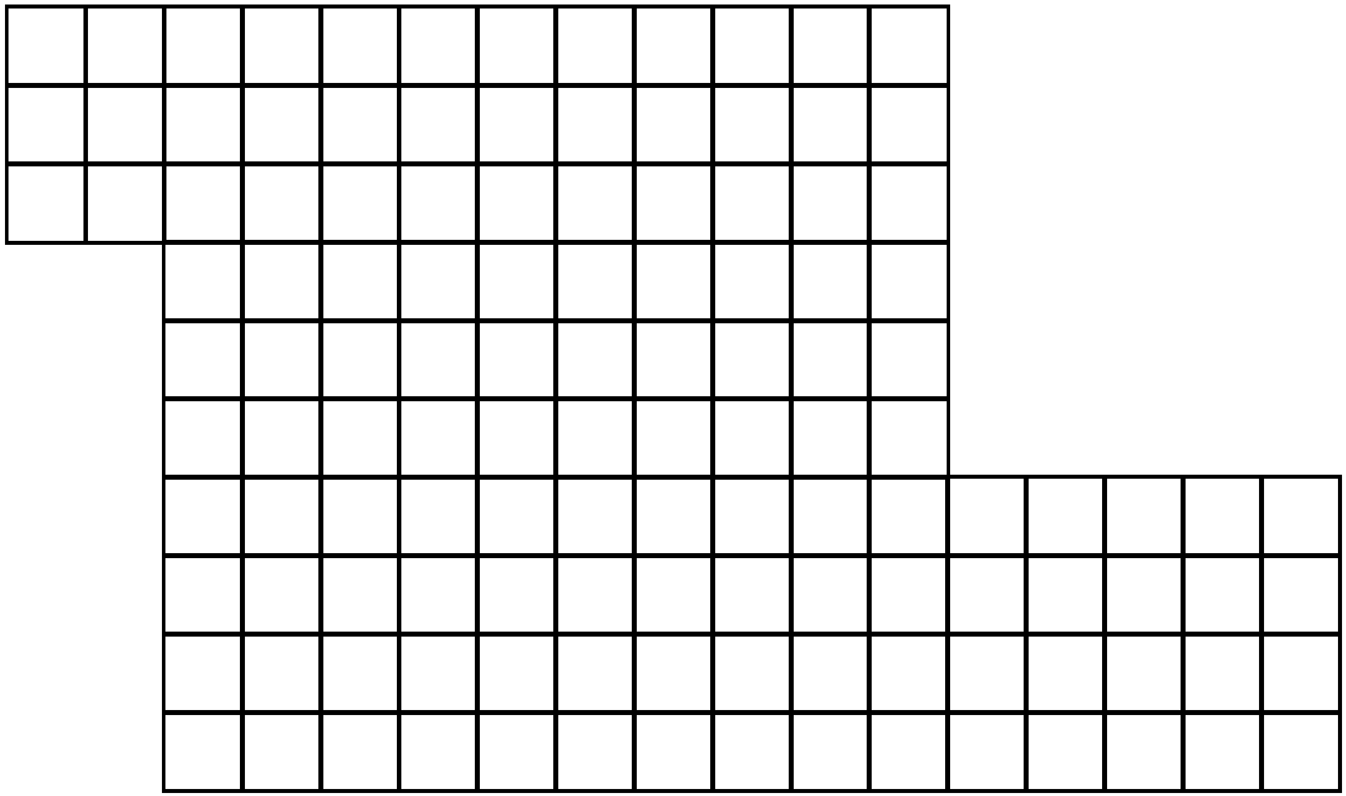
Chalmers University of Technology

SE-412 96 Göteborg, Sweden

# Example: A remediation site

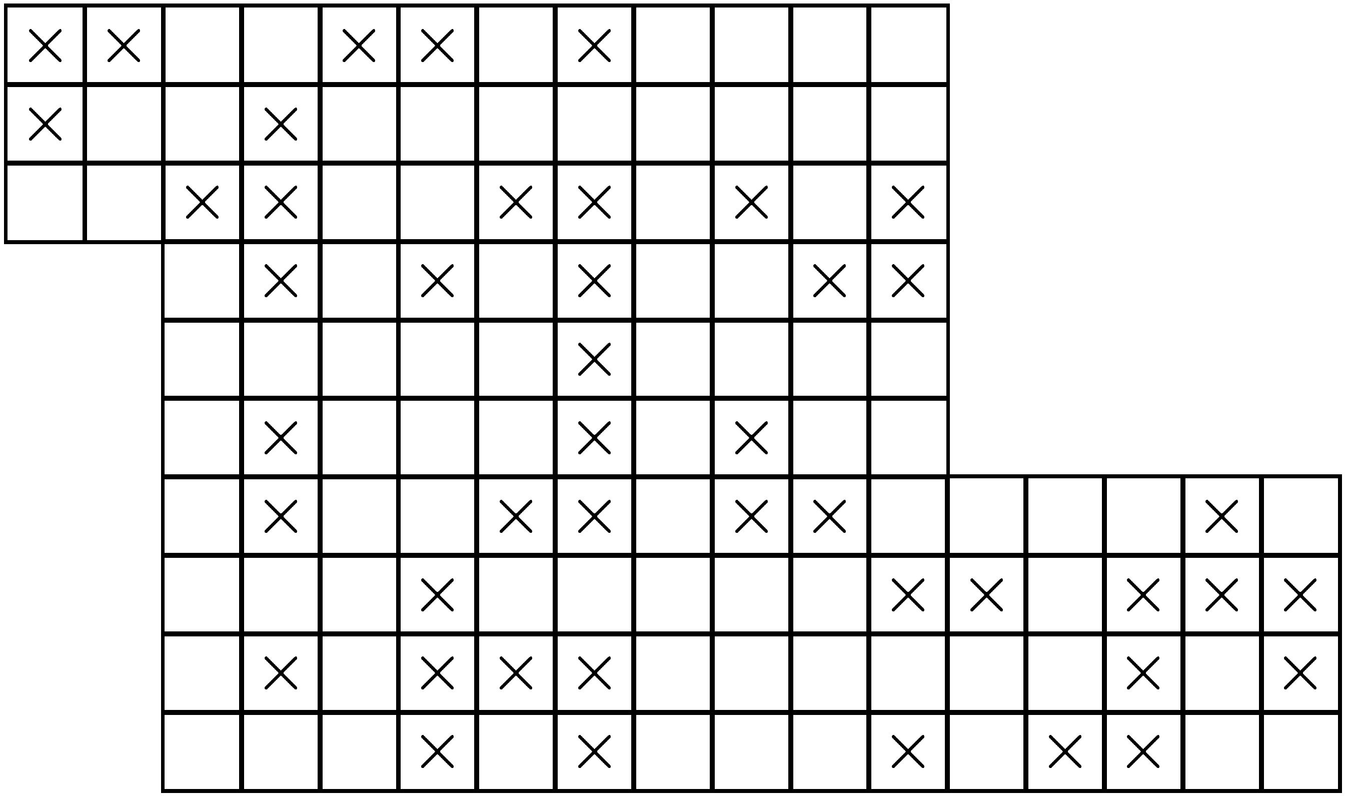


# Example: A remediation site



The site is partitioned into $N = 126$ Remediation Units.

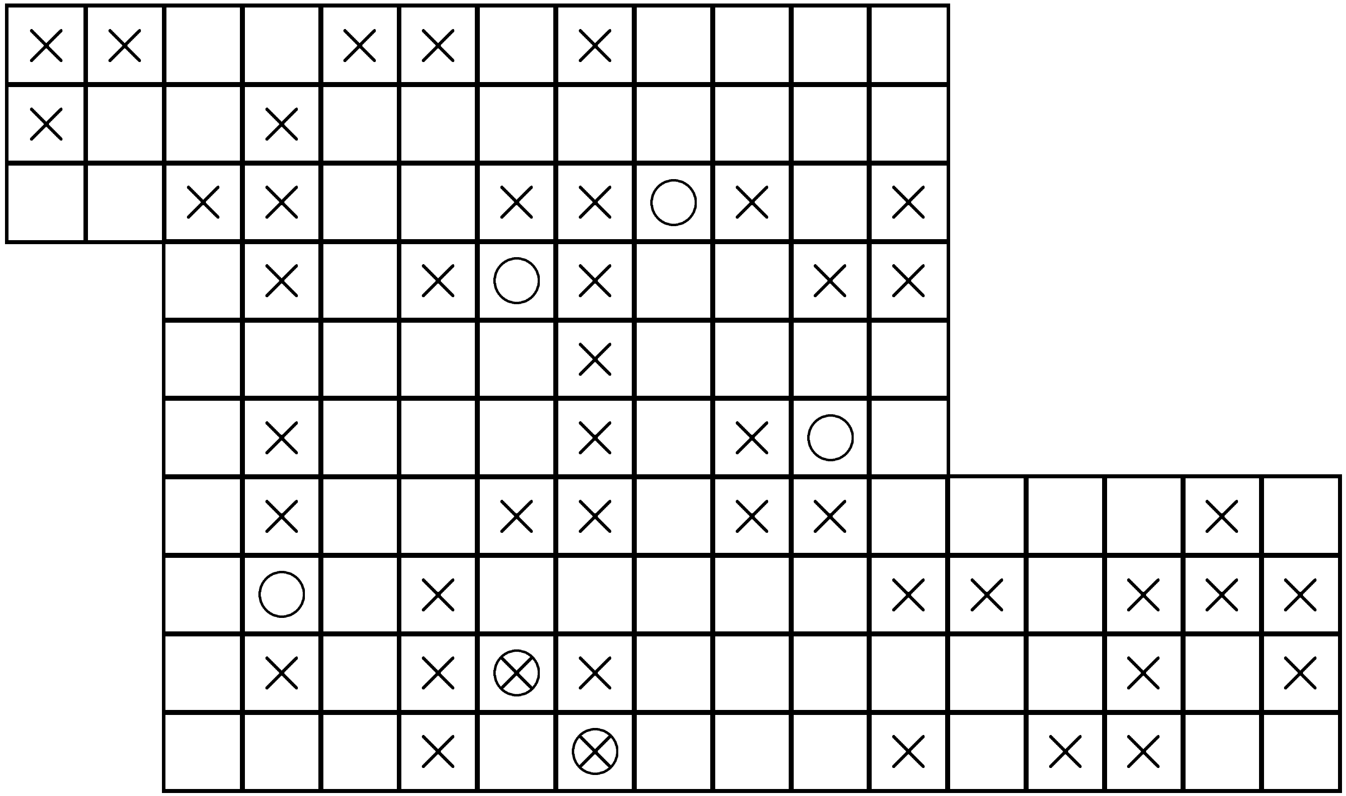
# Example: A remediation site



The site is partitioned into $N = 126$ Remediation Units.

We don't know that $t = 45$ RUs are contaminated.

# Example: A remediation site



The site is partitioned into $N = 126$ Remediation Units.

We don't know that $t = 45$ RUs are contaminated.

Of $n = 6$ sampled RUs, $x = 2$ turned out contaminated.

# Our task

is to remediate contaminated sites like this one in a cost-effective manner.

Alternatives:

0. Remediate all RUs.
1. Sample all RUs and remediate those that are contaminated.
2. Sample a fixed number of RUs and remediate the contaminated units found.
3. Sample a random number of RUs and remediate the contaminated units found.

# Notation

Denote by

- $t$ the number of contaminated RUs,
- $n$ the number of sampled RUs.

Let further

- $x$ be the number of contaminated RUs found by sampling

and put $y = n - x$.

# Risk measure, 1st choice

Assume that the prior risk is

$$(k_F + g_B)\mu_t$$

where $\mu_t$ is the mean of $t$ and $k_F$ is a failure cost, while $g_B$ denotes the benefits that are lost due to failure.

Denote by $k_R$ the remediation cost and let $k_M = c_M k_R$ be the cost of sampling one RU.

# Risk reduction

The mean risk reduction due to sampling of $n$ RUs is

$$(k_F + g_B)\left(\mu_t - E[\mu_{t-x|n,x}|n]\right) = (k_F + g_B)\mu_{x|n}$$

where $\mu_{x|n}$ is the mean number of contaminated RUs found.

The corresponding mean cost is

$$nk_M + \mu_{x|n}k_R = (k_F + g_B)(nc_M + \mu_{x|n})\alpha$$

where

$$\alpha = \frac{k_R}{k_F + g_B}$$

# Normalized risk reduction

The risk reduction per unit cost in Alternative 3 is

$$\frac{\mu_x}{(\mu_n c_M + \mu_x)\alpha}$$

# Normalized risk reduction

The *normalized risk reduction per unit cost* is

$$R_3 = \frac{\mu_x}{\mu_n c_M + \mu_x}$$

The corresponding quantity for Alternatives 2 and 1 is

$$R_2 = \frac{\mu_{x|n}}{n c_M + \mu_{x|n}}$$

$$R_1 = \frac{\mu_t}{N c_M + \mu_t}$$

# Prior knowledge

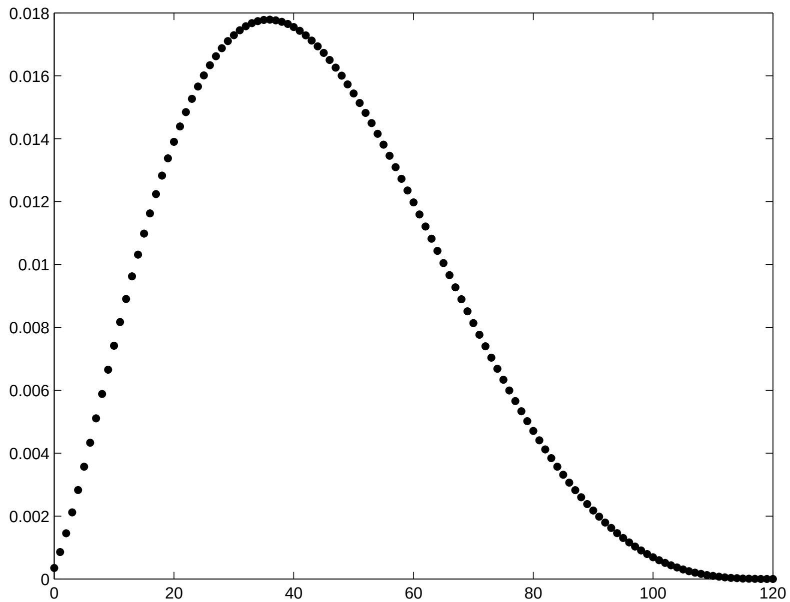
Our prior opinion on the number of contaminated RUs is modelled by a probability density.

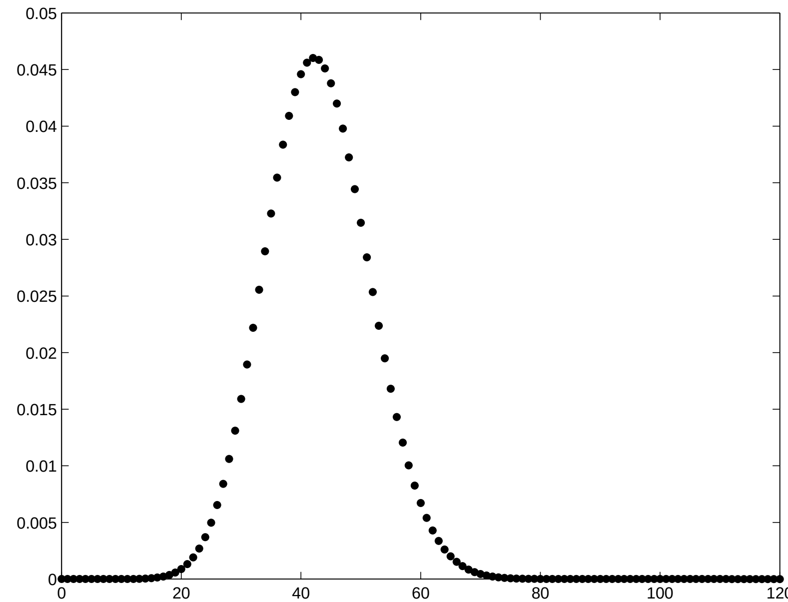
In this work, we only consider cases in which

$$t \sim \mathrm{BB}(N, a, b)$$

This is the Beta-Binomial distribution. The parameters $a, b > 0$ reflect our prior knowledge.

# Density of BB$(N, a, b)$



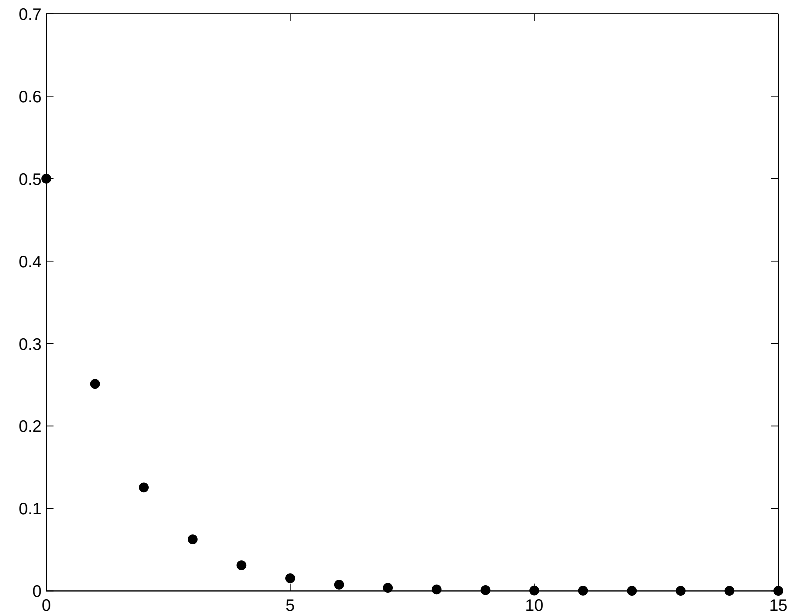




In both examples $N = 120$, and $\mu_t = 42.9$.

Left example has $a = 2.5, b = 4.5$.

Right example has $a = 25, b = 45$.

# A 3rd example




Here $N = 126, a = 1, b = 126$, and $P(t > 0) = 0.5$.

# Sampling

The only type of sampling considered is pure random sampling, i.e,

$$x|n,t \sim \text{HG}(N,n,t)$$

This is the Hyper-Geometric distribution.

Stratified sampling, sampling based on knowledge of the spatial distribution of the contaminant, etc, will not be considered.

# Sampling

The only type of sampling considered is pure random sampling, i.e,

$$x|n,t \sim \mathsf{HG}(N,n,t)$$

This is the Hyper-Geometric distribution.

Stratified sampling, sampling based on knowledge of the spatial distribution of the contaminant, etc, will not be considered.

By Bayes formula,

$$t-x|n,x \sim \mathsf{BB}(N-n,a+x,b+y)$$

# Example 1

$$N = 120,\ a = 2.5,\ b = 4.5$$
$$c_M = 0.1$$
prior risk $\propto \mu_t$

Sample RUs randomly until $\mu_{t-x|n,x} < 5$

Notice

$$\mu_{t-x|n,x} = (N-n)\,\frac{a+x}{a+b+n}$$

# Simulation 1

$$N = 120,\ a = 2.5,\ b = 4.5$$
$$c_M = 0.1$$
prior risk $\propto \mu_t$
Sample RUs randomly until $\mu_{t-x|n,x} < 5$

10 000 remediations were simulated, in which

$$t \sim \text{BB}(120, 2.5, 4.5)$$

Results:

- $R_3 = 0.7853$
- $R_1 = 0.7786$ (exact value is $0.7813$)
- $R_3/R_1 = 1.0087$

# 100 repetitions of Simulation 1

95 percent confidence intervals:

$$R_3 = 0.7885 \pm 0.0006$$
$$R_1 = 0.7811 \pm 0.0005$$
$$R_3/R_1 = 1.0095 \pm 0.0001$$

(recall $R_1 = 0.7813$).

Sign test of $H_0 : R_3 \leq R_1$ vs $H_1 : R_3 > R_1$ has $P$-value $0$.

# Simulation 2

$$N = 120,\ a = 2.5,\ b = 4.5$$
$$c_M = 0.1$$
prior risk $\propto \mu_t$
Sample RUs randomly until $\mu_{t-x|n,x} < 5$

10 000 remediations were simulated, in which

$$t \sim \text{BB}(120, 25, 45)$$

Results:

- $R_3 = 0.7836$
- $R_1 = 0.7813$ (exact value is $0.7813$)
- $R_3/R_1 = 1.0030$

# 100 repetitions of Simulation 2

95 percent confidence intervals:

$$R_3 = 0.7836 \pm 0.0003$$
$$R_1 = 0.7813 \pm 0.0003$$
$$R_3/R_1 = 1.0029 \pm 0.0001$$

(recall $R_1 = 0.7813$).

Sign test of $H_0 : R_3 \leq R_1$ vs $H_1 : R_3 > R_1$ has $P$-value $0$.

# Risk measure, 2nd choice

Assume that the total risk at the unremediated site is

$$(k_F + g_B)P(t > 0)$$

Then the mean risk reduction due to sampling is

$$(k_F + g_B)\,(P(t > 0) - P(t > x))$$

# Normalized risk reduction

The normalized risk reduction per unit cost in Alternatives 3, 2 and 1 is

$$R_3 = \frac{P(t > 0) - P(t > x)}{\mu_n c_M + \mu_x}$$

$$R_2 = \frac{P(t > 0) - P(t > x|n)}{n c_M + \mu_{x|n}}$$

$$R_1 = \frac{P(t > 0)}{N c_M + \mu_t}$$

# Simulation 3

$$N = 126,\ a = 1,\ b = 126$$
$$c_M = 0.1$$
$$\text{prior risk} \propto P(t > 0)$$
$$\text{Sample RUs randomly until } P(t > x|n, x) < 0.2$$

10 000 remediations were simulated, in which

$$t \sim \text{BB}(126, 1, 126)$$

Results:

- $R_3 = 0.0320$, $\mu_n = 87.2$
- $R_2|_{n=87.2} = 0.0279$ (exact value is $0.0281$)
- $R_3/R_2 = 1.1477$

# 100 repetitions of Simulation 3

95 percent confidence intervals:

$$R_3 = 0.0320 \pm 0.00002$$

$$R_2 = 0.0279 \pm 0.0000$$

$$R_3/R_2 = 1.1451 \pm 0.0005$$

Sign test of $H_0 : R_3 \leq R_2$ vs $H_1 : R_3 > R_2$ has $P$-value $0$.