Introduction to Rare Event Simulation

Brown University: Summer School on Rare Event Simulation

Jose Blanchet

Columbia University.
Department of Statistics,
Department of IEOR.
Goals / questions to address in this lecture

- The need for rare event simulation by introducing several problems.
Goals / questions to address in this lecture

- The need for rare event simulation by introducing several problems.
- Why do we care about rare event simulation?
Goals / questions to address in this lecture

- The need for rare event simulation by introducing several problems.
- Why do we care about rare event simulation?
- How does one measure the efficiency of rare event simulation algorithms?
Goals / questions to address in this lecture

- The need for rare event simulation by introducing several problems.
- Why do we care about rare event simulation?
- How does one measure the efficiency of rare event simulation algorithms?
- Describe basic rare event simulation techniques.
Why are Rare Events Difficult to Assess?

- Typically no closed forms
- But crudely implemented simulation might not be good
Why are Rare Events Difficult to Assess?

![Graph showing points in a 2D space with a red cluster and blue points.](image)
Why are Rare Events Difficult to Assess?

- Relative **mean squared error (MSE)** PER TRIAL = \( \frac{\text{stdev}}{\text{mean}} \)

\[
\frac{\sqrt{P(\text{RED})(1 - P(\text{RED}))}}{P(\text{RED})} \approx \frac{1}{\sqrt{P(\text{RED})}}.
\]

Moral of Story: Naive Monte Carlo needs \( N = \Theta \left( \frac{1}{P(\text{Rare Event})} \right) \) replications to obtain a given relative accuracy.
Why are Rare Events Difficult to Assess?

- Relative **mean squared error (MSE)** PER TRIAL = \( \frac{\text{stdev}}{\text{mean}} \)

\[
\frac{\sqrt{P(\text{RED})(1 - P(\text{RED}))}}{P(\text{RED})} \approx \frac{1}{\sqrt{P(\text{RED})}}.
\]

- **Moral of Story**: Naive Monte Carlo needs

\[
N = O \left( \frac{1}{P(\text{Rare Event})} \right)
\]

replications to obtain a given relative accuracy.
Rare event simulation = techniques for efficient estimation of rare-event probabilities AND conditional expectations given such rare events.
Rare event simulation = techniques for efficient estimation of rare-event probabilities AND conditional expectations given such rare events.

Arises in settings such as:
Rare event simulation = techniques for efficient estimation of rare-event probabilities AND conditional expectations given such rare events.

Arises in settings such as:

1. Finance and Insurance,
Rare event simulation = techniques for efficient estimation of rare-event probabilities AND conditional expectations given such rare events.

Arises in settings such as:

1. Finance and Insurance,
2. Queueing networks (operations and communications),
Rare event simulation = techniques for efficient estimation of rare-event probabilities AND conditional expectations given such rare events.

Arises in settings such as:

1. Finance and Insurance,
2. Queueing networks (operations and communications),
3. Environmental and physical sciences,
Rare event simulation = techniques for efficient estimation of rare-event probabilities AND conditional expectations given such rare events.

Arises in settings such as:

1. Finance and Insurance,
2. Queueing networks (operations and communications),
3. Environmental and physical sciences,
4. Statistical inference and combinatorics...
Introduction to Rare Event Simulation

- **Rare event simulation** = techniques for efficient estimation of rare-event probabilities AND conditional expectations given such rare events.
- Arises in settings such as:
  1. Finance and Insurance,
  2. Queueing networks (operations and communications),
  3. Environmental and physical sciences,
  4. Statistical inference and combinatorics...
  5. Stochastic optimization,
$V_i$’s i.i.d. (independent and identically distributed) claim sizes
Insurance Reserve Setup

- $V_i$’s i.i.d. (independent and identically distributed) claim sizes
- $\tau_i$’s i.i.d. inter-arrival times
Insurance Reserve Setup

- $V_i$’s i.i.d. (independent and identically distributed) claim sizes
- $\tau_i$’s i.i.d. inter-arrival times
- $A_n = $ time of the $n$-th arrival
Insurance Reserve Setup

- $V_i$’s i.i.d. (independent and identically distributed) claim sizes
- $\tau_i$’s i.i.d. inter-arrival times
- $A_n = \text{time of the } n\text{-th arrival}$
- Constant premium $p$
Insurance Reserve Setup

- $V_i$’s i.i.d. (independent and identically distributed) claim sizes
- $\tau_i$’s i.i.d. inter-arrival times
- $A_n = \text{time of the } n\text{-th arrival}$
- Constant premium $p$
- Reserve process

$$R(t) = b + pt - \sum_{j=1}^{N(t)} V_j$$
Insurance Reserve Setup

- $V_i$’s i.i.d. (independent and identically distributed) claim sizes
- $\tau_i$’s i.i.d. inter-arrival times
- $A_n = \text{time of the } n\text{-th arrival}$
- Constant premium $p$
- Reserve process
  \[ R(t) = b + pt - \sum_{j=1}^{N(t)} V_j \]
- $N(t) = \# \text{ arrivals up to time } t$
Insurance Reserve Plot

Plot of risk reserve

\[ R(t) \]

\[ b \]

\[ A_1 \quad A_2 \quad A_3 \quad t \]
Evaluating reserve at arrival times we get random walk.
Evaluating reserve at arrival times we get random walk

Suppose \( Y_1, Y_2, \ldots \) are i.i.d.

\[
S(n) = b + Y_1 + \ldots + Y_n
\]
Insurance Reserve and Random Walks

- Evaluating reserve at arrival times we get random walk.
- Suppose $Y_1, Y_2, \ldots$ are i.i.d.

$$S(n) = b + Y_1 + \ldots + Y_n$$

- $R(A_n) = S(n)$ reserve at arrival times with $Y_n = p\tau_n - V_n$.

Question:

What is the chance that $\inf_{t \geq 0} R(t) < 0$? And how does this happen?
Evaluating reserve at arrival times we get random walk

Suppose \( Y_1, Y_2, \ldots \) are i.i.d.

\[
S(n) = b + Y_1 + \ldots + Y_n
\]

\( R(A_n) = S(n) \) reserve at arrival times with \( Y_n = p\tau_n - V_n \).

**Question:** What is the chance that \( \inf \{ R(t) : t \geq 0 \} < 0 \)? And how does this happen?
How Does Ruin Occur with Light-Tailed Increments?

- $Y_i$'s are $N(0,1)$, $EY_i = 1$
How Does Ruin Occur with Light-Tailed Increments?

- $Y_i$'s are $N(0, 1)$, $EY_i = 1$
- Random walk conditioned on ruin

![Ruin with Gaussian Increments](image)
How Does Ruin Occur with Light-Tailed Increments?

- $Y_i$’s are $N(0,1)$, $EY_i = 1$
- Random walk conditioned on ruin
- Light tails: Exponential, Gamma, Gaussian, mixtures of these, etc.
How Does Ruin Occur with Light-Tailed Increments?

- $Y_i$’s are $N(0,1)$, $EY_i = 1$
- Random walk conditioned on ruin
- Light tails: Exponential, Gamma, Gaussian, mixtures of these, etc.
- Picture generated with Siegmund’s 76 algorithm
Y_i’s are t-distributed, EY_i = 1, Var(Y_i) = 1, 4 degrees of freedom.
Y_i’s are *t*-distributed, $EY_i = 1$, $Var(Y_i) = 1$, 4 degrees of freedom.

Heavy-tailed random walk conditioned on ruin.
• $Y_i$’s are $t$-distributed, $EY_i = 1$, $Var(Y_i) = 1$, 4 degrees of freedom.
• Heavy-tailed random walk conditioned on ruin.
• Picture generated with Blanchet and Glynn’s 09 algorithm.
A Two-node Tandem Network

- Poisson arrivals, exponential service times, and Markovian routing...
Questions of interest:

How would the system perform IF TOTAL POPULATION reaches $n$ inside a busy period?

How likely is this event under different parameters / designs?
Naive Benchmark: Crude Monte Carlo

• Say $\lambda = .1$, $\mu_1 = .5$, $\mu_2 = .4$, $p = .1$
Naive Benchmark: Crude Monte Carlo

- Say $\lambda = .1$, $\mu_1 = .5$, $\mu_2 = .4$, $p = .1$
- Overflow probability $\approx 10^{-7}$
Naive Benchmark: Crude Monte Carlo

- Say $\lambda = .1$, $\mu_1 = .5$, $\mu_2 = .4$, $p = .1$
- Overflow probability $\approx 10^{-7}$
- Each replication in crude Monte Carlo $\approx 10^{-3}$ secs
Naive Benchmark: Crude Monte Carlo

Say $\lambda = .1$, $\mu_1 = .5$, $\mu_2 = .4$, $\rho = .1$

Overflow probability $\approx 10^{-7}$

Each replication in crude Monte Carlo $\approx 10^{-3}$ secs

Time to estimate with 10% precision $\approx \frac{10^7}{1000} = 10^4$ secs $\approx 2.7$ hours
Say $\lambda = .1$, $\mu_1 = .5$, $\mu_2 = .4$, $p = .1$

Overflow probability $\approx 10^{-7}$

Each replication in crude Monte Carlo $\approx 10^{-3}$ secs

Time to estimate with 10% precision $\approx 10^7/1000 = 10^4$ secs $\approx 2.7$ hours

What if you want to do sensitivity analysis for a wide range of parameters?
Naive Benchmark: Crude Monte Carlo

- Say $\lambda = .1$, $\mu_1 = .5$, $\mu_2 = .4$, $p = .1$
- Overflow probability $\approx 10^{-7}$
- Each replication in crude Monte Carlo $\approx 10^{-3}$ secs
- Time to estimate with 10% precision $\approx 10^7 / 1000 = 10^4$ secs $\approx 2.7$ hours
- What if you want to do sensitivity analysis for a wide range of parameters?
- What about different network designs?
Question: How does the TOTAL CONTENT OF THE SYSTEM is most likely to reach a give level in an operation cycle?
Question: How does the TOTAL CONTENT OF THE SYSTEM is most likely to reach a give level in an operation cycle?

Say $\lambda = .2$, $\mu_1 = .3$, $\mu_2 = .5$... How can total content reach 50?
Question: How does the TOTAL CONTENT OF THE SYSTEM is most likely to reach a given level in an operation cycle?
Say $\lambda = .2$, $\mu_1 = .3$, $\mu_2 = .5...$ How can total content reach 50?
Naive Monte Carlo takes $\approx 115$ days
Question: How does the TOTAL CONTENT OF THE SYSTEM is most likely to reach a given level in an operation cycle?

Say $\lambda = .2, \mu_1 = .3, \mu_2 = .5...$ How can total content reach $50$?

Naive Monte Carlo takes $\approx 115$ days

Each picture below took $\approx .01$ seconds
$\lambda = 0.1, \mu_1 = 0.3, \mu_2 = 0.5$
\( \lambda = .1, \mu_1 = .5, \mu_2 = .2 \)
$\lambda = 1, \mu_1 = 0.4, \mu_2 = 0.4$
Assume the estimator is unbiased.
Several Notions of Efficiency

- Assume the estimator is unbiased.
- $E(Z) = P(A)$ and we assume $P(A) \rightarrow 0$. 

Assume the estimator is unbiased.

\[ E(Z) = P(A) \] and we assume \( P(A) \rightarrow 0. \)

**Logarithmic Efficiency (weakly efficiency or asymptotic optimality):**

\[
\lim_{P(A) \to 0} \frac{\log E(Z^2)}{2 \log P(A)} = 1.
\]
Assume the estimator is unbiased.

\[ E(Z) = P(A) \] and we assume \( P(A) \to 0 \).

**Logarithmic Efficiency (weakly efficiency or asymptotic optimality):**

\[
\lim_{P(A)\to0} \frac{\log E(Z^2)}{2 \log P(A)} = 1.
\]

**Bounded Relative Error**

\[
\frac{E(Z^2)}{P(A)^2} = O(1) \quad \text{as} \quad P(A) \to 0.
\]
If $Z = I(A)$ then $E(Z^2) = E(Z) = P(A)$
If $Z = I(A)$ then $E(Z^2) = E(Z) = P(A)$

Therefore

$$\lim_{P(A) \to 0} \frac{\log P(A)}{2 \log P(A)} = \frac{1}{2}.$$
If \( Z = I(A) \) then \( E(Z^2) = E(Z) = P(A) \)

Therefore

\[
\lim_{P(A) \to 0} \frac{\log P(A)}{2 \log P(A)} = \frac{1}{2}.
\]

How to design efficient rare event simulation estimators? Importance Sampling and other variance reduction techniques...
Importance sampling (I.S.): sample from the important region and correct via likelihood ratio

RED AREA \approx \text{PROPORTION DARTS IN RED AREA} \times \frac{1}{9}
Goal: Estimate $P(A) > 0$
Goal: Estimate $P(A) > 0$

Choose $\tilde{P}$ and simulate $\omega$ from it
Importance Sampling

- **Goal:** Estimate $P(A) > 0$
- Choose $\hat{P}$ and simulate $\omega$ from it
- **Importance Sampling (I.S.)** estimator *per trial* is

$$I.S.\text{Estimator} = L(\omega) I(\omega \in A),$$

where $L(\omega)$ is the likelihood ratio (i.e. $L(\omega) = \frac{dP(\omega)}{d\hat{P}(\omega)}$).
Importance Sampling

- **Goal:** Estimate \( P(A) > 0 \)
- Choose \( \tilde{P} \) and simulate \( \omega \) from it
- **Importance Sampling (I.S.)** estimator *per trial* is

\[
I.S.\text{Estimator} = L(\omega) I(\omega \in A),
\]

where \( L(\omega) \) is the likelihood ratio (i.e. \( L(\omega) = dP(\omega) / d\tilde{P}(\omega) \)).

- **NOTE:** \( \tilde{P}(\cdot) \) is called a change-of-measure
Suppose we choose $\tilde{P}(\cdot) = P(\cdot|A)$

$$L(\omega) = \frac{dP(\omega)}{I(\omega \in A) dP(\omega) / P(A)} I(\omega \in A) = P(A)$$
Importance Sampling

Suppose we choose $\tilde{P}(\cdot) = P(\cdot | A)$

$$L(\omega) = \frac{dP(\omega)}{I(\omega \in A) dP(\omega) / P(A)} I(\omega \in A) = P(A)$$

Estimator has zero variance, but requires knowledge of $P(A)$
Suppose we choose \( \tilde{P}(\cdot) = P(\cdot \mid A) \)

\[
L(\omega) = \frac{dP(\omega)}{I(\omega \in A) \frac{dP(\omega)}{P(A)} I(\omega \in A)} = P(A)
\]

Estimator has zero variance, but requires knowledge of \( P(A) \)

**Lesson:** Try choosing \( \tilde{P}(\cdot) \) close to \( P(\cdot \mid A) \)!
Suppose $X_1, X_2, \ldots$ are i.i.d. and $H(\theta) = \log E \exp(\theta X) < \infty$ for $\theta$ in a neighborhood of the origin.
Suppose $X_1, X_2, \ldots$ are i.i.d. and $H(\theta) = \log E \exp(\theta X) < \infty$ for $\theta$ in a neighborhood of the origin.

Assume that $E(X_i) = 0$. 
Suppose $X_1, X_2, \ldots$ are i.i.d. and $H(\theta) = \log E \exp(\theta X) < \infty$ for $\theta$ in a neighborhood of the origin.

Assume that $E(X_i) = 0$.

Suppose that for each $a > 0$, there is $\theta_a$ such that $H'(\theta_a) = a$. 

Consider the problem of estimating via simulation $P(S(n) > na)$, where $S(n) = X_1 + \ldots + X_n$. 

Blanchet (Columbia)
Suppose $X_1, X_2, \ldots$ are i.i.d. and $H(\theta) = \log E \exp(\theta X) < \infty$ for $\theta$ in a neighborhood of the origin.

Assume that $E(X_i) = 0$.

Suppose that for each $a > 0$, there is $\theta_a$ such that $H'(\theta_a) = a$.

Consider the problem of estimating via simulation

$$P(S(n) > na),$$

where $S(n) = X_1 + \ldots + X_n$. 
Suppose that $X_i$ has density $f(\cdot)$ and define

$$f_\theta(x) = \exp(\theta x - H(\theta)) f(x).$$
Suppose that $X_i$ has density $f(\cdot)$ and define

$$f_\theta(x) = \exp(\theta x - H(\theta)) f(x).$$

Note that

$$\log E\theta(\exp(\eta X)) = \log \int f_\theta(x) \exp(\eta x) \, dx = \log \int f(x) \exp((\eta + \theta)x - H(\theta)) \, dx = H(\eta + \theta) - H(\theta).$$
Exponential Tilting

- Suppose that $X_i$ has density $f(\cdot)$ and define
  \[
  f_\theta (x) = \exp(\theta x - H(\theta)) f(x).
  \]

- Note that
  \[
  \log E_\theta (\exp(\eta X)) = \log \int f_\theta (x) \exp(\eta x) \, dx
  \]
  \[
  = \log \int f(x) \exp((\eta + \theta) x - H(\theta)) \, dx
  \]
  \[
  = H(\eta + \theta) - H(\theta).
  \]

- Therefore $E_\theta (X) = H'(\theta)$ and $\text{Var}_\theta (X) = H''(\theta)$. 
Cramer’s theorem (strong version): (Sketch)

\[ P(S_n > na) = \int f(x_1) \ldots f(x_n) I(s_n > na) \, dx_{1:n} \]

\[ = \int_{\{s_n > na\}} f_{\theta_a}(x_1) \ldots f_{\theta_a}(x_n) \exp\left(-\theta_a s_n + nH(\theta_a)\right) \, dx_{1:n} \]

\[ = \exp\left(-n[a\theta_a - H(\theta_a)]\right) E_{\theta_a}\left[\exp\left(-\theta_a (S_n - na)\right) I(S_n > na)\right] \]

\[ = \exp\left(-nI(a)\right) E_{\theta_a}\left[\exp\left(-\theta_a (S_n - na)\right) I(S_n > na)\right]. \]
Cramer’s theorem (strong version): (Sketch)

\[
P(S_n > na) = \int f(x_1) \ldots f(x_n) I(s_n > na) \, dx_{1:n}
\]

\[
= \int_{\{s_n > na\}} f_{\theta_a}(x_1) \ldots f_{\theta_a}(x_n) \exp(-\theta_a s_n + nH(\theta_a)) \, dx_{1:n}
\]

\[
= \exp(-n[a\theta_a - H(\theta_a)]) E_{\theta_a}[\exp(-\theta_a (S_n - na)) I(S_n > na)]
\]

\[
= \exp(-nl(a)) E_{\theta_a}[\exp(-\theta_a (S_n - na)) I(S_n > na)].
\]

Use CLT and approximate \((S_n - na)\) by \(N(0, nH''(\theta_a))\) to obtain

\[
E_{\theta_a}[\exp(-\theta_a (S_n - na)) I(S_n > na)] \approx \frac{1}{n^{1/2}(2\pi H''(\theta_a))^{1/2} \theta_a}.
\]
Why is Exponential tilting Useful?

Theorem

Under the assumptions imposed

\[ P (X_1 \in dx_1 | S_n > na) \to P_{\theta_a} (X_1 \in dx_1). \]

- In simple words, exponential tilting approximates the zero-variance (optimal!) change-of-measure!

Proof.

\[ P (X_1 \in dx_1 | S_n > na) = \frac{P (S_{n-1} > na - x_1)}{P (S_n > na)} f (x_1) \, dx. \]

Now apply Cramer’s theorem with \( a \leftarrow (na - x_1) / (n - 1) \) and \( n \leftarrow n - 1 \) in the numerator and simplify, observing that

\[ \theta \left( \frac{na - x_1}{n - 1} \right) = \theta \left( a + \frac{a - x_1}{n - 1} \right) \approx \theta (a) + \frac{1}{H''(\theta_a)} \frac{a - x_1}{n - 1}. \]
First Asymptotically Optimal Rare Event Simulation Estimator

- Recall the identity used in Cramer’s theorem

\[
P(S_n > na) = \int f(x_1) \ldots f(x_n) I(s_n > na) \, dx_{1:n} \\
= \int_{\{s_n > na\}} f_{\theta_a}(x_1) \ldots f_{\theta_a}(x_n) \exp(-\theta_a s_n + nH(\theta_a)) \, dx_{1:n} \\
= E_{\theta_a}[\exp(-\theta_a S_n + nH(\theta_a)) \, I(S_n > na)].
\]
Recall the identity used in Cramer’s theorem

\[
P(S_n > na) = \int f(x_1) \cdots f(x_n) I(s_n > na) \, dx_1 \cdots dx_n
\]

\[
= \int_{\{s_n > na\}} f_{\theta_a}(x_1) \cdots f_{\theta_a}(x_n) \exp(-\theta_a s_n + nH(\theta_a)) \, dx_1 \cdots dx_n
\]

\[
= E_{\theta_a} \left[ \exp(-\theta_a S_n + nH(\theta_a)) I(S_n > na) \right].
\]

**IMPORTANCE SAMPLING ESTIMATOR:**

\[
Z = \exp(-\theta_a S_n + nH(\theta_a)) I(S_n > na),
\]

with \(X_1, \ldots, X_n\) simulated using the density \(f_{\theta_a}(\cdot)\).
First Asymptotically Optimal Rare Event Simulation Estimator

- Recall the identity used in Cramer’s theorem

\[ P (S_n > na) = \int f(x_1) \ldots f(x_n) I(s_n > na) \, dx_{1:n} \]

\[ = \int_{\{s_n > na\}} f_{\theta_a}(x_1) \ldots f_{\theta_a}(x_n) \exp(-\theta_a s_n + nH(\theta_a)) \, dx_{1:n} \]

\[ = E_{\theta_a} [\exp(-\theta_a S_n + nH(\theta_a)) I(S_n > na)] . \]

- IMPORTANCE SAMPLING ESTIMATOR:

\[ Z = \exp(-\theta_a S_n + nH(\theta_a)) I(S_n > na) , \]

with \( X_1, \ldots, X_n \) simulated using the density \( f_{\theta_a} (\cdot) \).

- Let us verify now that \( Z \) is asymptotically efficient!
IMPORTANCE SAMPLING ESTIMATOR:

\[ Z = \exp(-\theta_a S_n + nH(\theta_a)) I(S_n > na), \]

with \( X_1, ..., X_n \) simulated using the density \( f_{\theta_a}(\cdot) \).
First Asymptotically Optimal Rare Event Simulation Estimator

- **IMPORTANCE SAMPLING ESTIMATOR:**

\[ Z = \exp(-\theta_a S_n + nH(\theta_a)) \mathbb{I}(S_n > na), \]

with \( X_1, \ldots, X_n \) simulated using the density \( f_{\theta_a}(\cdot) \).

- Therefore,

\[
E_{\theta_a}(Z^2) = E_{\theta_a} \exp(-2\theta_a S_n + 2nH(\theta_a)) \mathbb{I}(S_n > na) \\
= \exp(-2nI(a)) E_{\theta_a} [\exp(-2\theta_a(S_n - na)) \mathbb{I}(S_n > na)].
\]
IMPORTANCE SAMPLING ESTIMATOR:

\[ Z = \exp(-\theta_a S_n + nH(\theta_a)) I(S_n > na), \]

with \( X_1, \ldots, X_n \) simulated using the density \( f_{\theta_a} (\cdot) \).

Therefore,

\[
E_{\theta_a} (Z^2) = E_{\theta_a} \exp(-2\theta_a S_n + 2nH(\theta_a)) I(S_n > na)
= \exp(-2nl(a)) E_{\theta_a} [\exp(-2\theta_a(S_n - na)) I(S_n > na)].
\]

Thus

\[
\frac{\log E_{\theta_a} (Z^2)}{2 \log P(S_n > na)} = \frac{2nl(a)}{2nl(a)} + o(1) \to 1.
\]
How to Simulate under Exponential Tilting

- **Key is to identify:** \( H_\theta (\eta) = H (\eta + \theta) - H (\theta) \).
Key is to identify: $H_\theta(\eta) = H(\eta + \theta) - H(\theta)$.

Example 1: Gaussian if $X \sim N(\mu, 1)$

$$H(\theta) = \frac{\theta^2}{2} + \mu \theta, \quad H_\theta(\eta) = \frac{\eta^2}{2} + (\theta + \mu) \eta.$$  

Thus, under $f_\theta$, $X \sim N(\mu, 1)$. 

Key is to identify: $H_\theta (\eta) = H (\eta + \theta) - H (\theta)$.

Example 1: Gaussian if $X \sim N (\mu, 1)$

$$H (\theta) = \theta^2 / 2 + \mu \theta, \quad H_\theta (\eta) = \eta^2 / 2 + (\theta + \mu) \eta.$$ 

Thus, under $f_\theta$, $X \sim N (\mu, 1)$.

Example 2: If $X \sim Poisson (\lambda)$, then

$$H (\theta) = \lambda (\exp (\theta) - 1), \quad H_\theta (\eta) = \lambda \exp (\theta) \cdot (\exp (\eta) - 1).$$

Thus, under $f_\theta$, $X \sim Poisson (\lambda \exp (\theta))$. 

Blanchet (Columbia)
**Conclusions**

1. Rare event simulation is a powerful numerical tool applicable in a wide range of examples.
Conclusions

1. Rare event simulation is a powerful numerical tool applicable in a wide range of examples.

2. Importance sampling can be used to construct efficient estimators.
Conclusions

1. Rare event simulation is a powerful numerical tool applicable in a wide range of examples.
2. Importance sampling can be used to construct efficient estimators.
3. Best importance sampler is the conditional distribution given the event of interest.
Conclusions

1. Rare event simulation is a powerful numerical tool applicable in a wide range of examples.

2. Importance sampling can be used to construct efficient estimators.

3. Best importance sampler is the conditional distribution given the event of interest.

4. Efficient algorithms come basically in two flavors: logarithmically and strongly efficient.
Conclusions

1. Rare event simulation is a power numerical tool applicable in a wide range of examples.
2. Importance sampling can be used to construct efficient estimators.
3. Best importance sampler is the conditional distribution given the event of interest.
4. Efficient algorithms come basically in two flavors: logarithmically and strongly efficient.
5. Exponential tilting can be used to approximate the conditional distribution given the event of interest.