LARGE DEVIATIONS

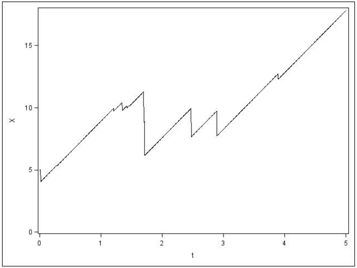
Jeffrey Collamore
University of Copenhagen

Historical motivation: Ruin problem

Lundberg (1903). Cramér-Lundberg process:

$$X_t = u + ct - \sum_{i=1}^{N_t} \zeta_i$$

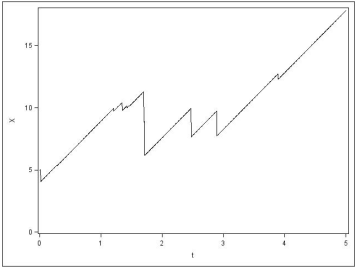
X_t =capital of insurance company at time t,
c =premiums income,
$\{\zeta_i\}$ claims losses.



Ruin problem (cont.)

Let $\Psi(u) = \mathbf{P}\left\{X_t < 0, \text{ for some } t \geq 0 \middle| X_0 = u\right\}$.
Lundberg (1903), Cramér (1930):

$$\Psi(u) \sim Ce^{-Ru} \quad \text{as } u \to \infty.$$



Ruin problem (cont.)

Why is this a large deviations problem?

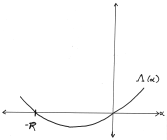
First, exponential decay as $u \to \infty$; in particular,

$$\lim_{u \to \infty} \frac{1}{u} \log \mathbf{P}\{\text{ruin}\} = -R,$$

where R is the *large deviation rate function*.

Second, proof is based on a change-of-measure argument.
Shifted measure: If $X_1 \sim \mu$, then consider

$$\mu_R(dx) := e^{-Rx} \mu(dx).$$



Statistics: sample means

$X_1, X_2, \ldots$ i.i.d. $\subset \mathbb{R}$.
Mean $\mathbf{E}[X_1] = m$.
Sample mean:

$$\bar{X}_n := \frac{X_1 + \cdots + X_n}{n}.$$

Law of large numbers:

$$\mathbf{P}\left\{|\bar{X}_n - m|\right\} \to 0 \quad \text{as} \quad n \to \infty.$$

How fast?
E.g., what is

$$\lim_{n\to\infty} \mathbf{P}\left\{\bar{X}_n > a\right\} \quad \text{for} \quad a > m?$$

Statistics: sample means

Central limit approximation:
Let $S_n = X_1 + \cdots + X_n$.
Then by the central limit theorem,

$$\bar{X}_n = \frac{S_n}{n} \approx \mathrm{Normal}\left(m, \frac{\sigma^2}{n}\right).$$

How good is this estimate?

Example

Suppose $\{X_i\}$ is i.i.d. Bernoulli sequence; thus,

$$X_i = \begin{cases} 1 & \text{w.p. } 1/2, \\ 0 & \text{w.p. } 1/2. \end{cases}$$

Then $S_n = X_1 + \cdots + X_n \sim$ binomial.
Now according to the central limit theorem, since $m = \sigma = 1/2$,

$$\begin{aligned} \mathbf{P}\left\{\frac{S_n}{n} > a\right\} &= \mathbf{P}\left\{\frac{S_n - n \cdot \frac{1}{2}}{\frac{1}{2}\sqrt{n}} > \sqrt{n}(2a-1)\right\} \\ &\approx 1 - \Phi\left(\sqrt{n}(2a-1)\right). \end{aligned}$$

But if $a > 1$, this probability is actually $= 0$.

Cramér's classical estimate

Assume finite moment generating fn.: $\lambda(\alpha) := \mathbf{E}\left[e^{\alpha X_i}\right] < \infty$.

Theorem (Cramér, 1938)

For any $a > m$,

$$\mathbf{P}\left\{\frac{S_n}{n} > a\right\} \sim \frac{c}{\sqrt{n}} e^{-nI(a)} \quad \text{as} \quad n \to \infty.$$

Shows:

- Exponential decay.
- Decay according to a "rate function" $I(x)$ (the convex conjugate of $\log \lambda$). This rate function will be *different* for every μ, where $X_i \sim \mu$.

Empirical measures

$\{X_i\}$ i.i.d., $X_i : \Omega \to \mathbb{S} = \{a_1, \ldots, a_d\}$.
Empirical measure:

$$L_n(a_j) = \frac{1}{j} \sum_{i=1}^{n} \delta_{X_i}(a_j),$$

that is, the proportion of observed samples which $= a_j$.
Note: L_n is a probability measure on $\mathbb{S}$.

Question: How likely will we empirically observe an $L_n(\cdot)$,
given the true distribution function.

Empirical measures (cont.)

Sanov (1957): If $\mathcal{S}$ is a set of empirical measures, then

$$\mathbf{P}\left\{L_n \in \mathcal{S}\right\} \approx e^{-nI(\mathcal{S})};$$

more precisely,

$$\lim_{n\to\infty} \log \mathbf{P}\left\{L_n \in \mathcal{S}\right\} = -I(\mathcal{S}).$$

Related problem: Brownian motion in a “tube" (Shilder, 1966). Consider

$$\mathbf{P}\left\{B_n(t) \in \Gamma(t) : 0 \leq t \leq 1\right\}, \quad n \to \infty,$$

where $B_n(t) = \frac{1}{\sqrt{n}} B(t)$ and $B(t)$ is Brownian motion.

References

DEMBO, A. and ZEITOUNI, O. (1998). Large Deviations Techniques and Applications. Springer, New York (2nd ed.).

VARADHAN, S.R.S. (1984). Large Deviations and Applications. SIAM, Philadelphia.

ELLIS, R.S. (1985). Entropy, Large Deviations and Statistical Mechanics. Springer, New York.

FREIDLIN, M.I. and WENTZELL, A.D. (1984). Random Perturbations of Dynamical Systems. Springer, New York.

COLLAMORE, J.F. and VIDYASHANKAR, A.N. Modern Large Deviations Theory: Methods and Applications. In preparation.