

Monte Carlo, Fall, 2005

Homework 1, due November 8

Problem 1. Suppose X is a random variable that takes value x_j with probability f_j , with $1 \leq j \leq s$. Let X_k , $1 \leq k \leq n$ be independent samples of X and let $N_j = \#\{X_k = x_j\}$ be the resulting hit counts. The *empirical probabilities* are $R_j = N_j/n$ and are probabilities ($R_j \geq 0$ for all j , and $\sum_j R_j = 1$). Write $\vec{R} = (R_1, \dots, R_s)$. The multinomial coefficient is (with $n_j \geq 0$ for all j , and $\sum_j n_j = n$)

$$C(n, n_1, \dots, n_s) = C(n, \vec{n}) = n! / \prod_{j=1}^s n_j! .$$

Then

$$P(\vec{N} = \vec{n}) = C(n, \vec{n}) \prod_{j=1}^s f_j^{n_j} . \quad (1)$$

(a) Use Stirling's formula,

$$n! \approx n^n e^{-n} ,$$

to write an approximation to (1) of the form

$$P(\vec{n}) \approx \frac{1}{Z} e^{-n\phi(r_1, \dots, r_n)} ,$$

where

$$r_j = n_j/n .$$

(b) Find the minimum of ϕ and the Taylor series about this minimum to second order to derive a Gaussian approximation to the distributin of \vec{N} for large n .

(c) Get the result of (b) above using the vector central limit theorem in R^s .

(d) Show that

$$P(\vec{R} = \vec{r}) \approx \frac{1}{Z} e^{-n\psi(\vec{r}, \vec{f})} ,$$

and find the formula for ψ . This is the *relative entropy* of \vec{r} with respect to \vec{f} .

(e) Solve the optimization problem

$$\min_{\vec{r}} \psi(\vec{r}, \vec{f}) \quad \text{subject to} \quad \sum_{j=1}^s x_j r_j = b .$$

Show that the optimal \vec{r} has the form

$$r_j = \frac{1}{Z(\lambda)} e^{\lambda x_j} f_j .$$

(Hint: Use Lagrange multipliers to inforce the two constraints.)

- (f) Show that if $b > E_f[X]$, the above implies the Cramer large deviation formula for discrete random variables.

Problem 2. Let $B \subset R^3$ be the unit ball of points y with $|y| \leq 1$. Let $h_0(y)$ be the uniform probability density inside B .

- (a) Write a program that samples h_0 by rejection from the uniform density in the smallest cube that contains B . Use histograms to check that $R = |Y|$ and $W = (Y_1 + 2Y_2 + 3Y_3)/\sqrt{14}$ have the correct one dimensional densities. Use small enough bins and large enough sample sizes so that the agreement between the empirical histogram and the expected bin counts is very good.
- (b) Calculate $Z(\lambda) = E[e^{\lambda Y_3}]$. Use the histogram from (a) to give a Monte Carlo estimate of $Z(\lambda)$ in the range $0 \leq \lambda \leq 6$. Give error bars and comment on the relative accuracy of the estimate for large λ .

- (c) If

$$h_\lambda(y) = \frac{1}{Z(\lambda)} e^{\lambda y_3} h_0(y),$$

write a program to sample the one dimensional random variable Y_3 by rejection from a piecewise linear trial distribution. Use a histogram to verify that your sampler is correct.

- (d) Write a program that samples the conditional density of (Y_1, Y_2) given the Y_3 sample from (c). Note that this does not depend on λ . Check this is correct by making a histogram of W with $\lambda = 0$.
- (e) Take $n = 100$ and

$$S = \sum_{k=1}^n Y_{3,k},$$

the sum of the third components of n $Y \in B$. Take a large number of samples of S when $Y \sim h_0$ and plot

$$\psi_n(b) = \frac{1}{n} \ln(P(S \geq nb)) \tag{2}$$

as a function of b (with error bars of some kind). Comment on the accuracy of your Monte Carlo estimator as b increases.

- (f) Now create a histogram with L weighted samples, S_k, L_k . If sample S_k lands in bin j , it adds weight L_k to the total weight of bin j . For each k , choose an independent $\lambda_k \in [0, 6]$ uniformly distributed, then choose $Y_k \sim h_{\lambda_k}$ and calculate the likelihood ratio L_k . Estimate (2) and comment on the differing accuracy for a range of b .