Tutorial and Lab Problems # 5
MATH3871/MATH5970

1. **Properties of Importance Sampling.** Suppose we wish to estimate $\alpha := \mathbb{E}_\pi \phi(X) = \int \pi(x)\phi(x)\mathrm{d}x$ using the importance sampling estimator:

$$\widehat{\alpha}_t := \frac{1}{t}\sum_{k=1}^{t} \overbrace{\frac{\pi(X_k)}{g(X_k)}}^{W_k} \phi(X_k), \qquad X_1, \ldots, X_t \sim_{\text{iid}} g(x)$$

Show that the resulting estimator is unbiased and has variance $\mathbb{V}\text{ar}(\widehat{\alpha}_t) = \frac{\mathbb{E}Z^2}{t}$, where $Z_k := W_k\phi(X_k) - \alpha$.

2. **Ratio Estimator.** Suppose that $(X_1, Y_1), \ldots, (X_n, Y_n)$ are iid copies of $\boldsymbol{X} = (X, Y)^\top$, where $\mathbb{E}\boldsymbol{X} = (\mu_X, \mu_Y)^\top$ and $\mathbb{C}\text{ov}(\boldsymbol{X}) = \Sigma$. Use the delta approximation (see below for reference) to show that

$$\sqrt{n}\left(\frac{\bar{X}_n}{\bar{Y}_n} - \alpha\right) \xrightarrow{\mathrm{d}} \mathsf{N}(0, \sigma^2),$$

where $\alpha := \mu_X/\mu_Y$ and

$$\sigma^2 := \frac{\mathbb{V}\text{ar}(X) - 2\alpha\mathbb{C}\text{ov}(X, Y) + \alpha^2\mathbb{V}\text{ar}(Y)}{\mu_Y^2}.$$

Hint: in the delta method use $g(x, y) = x/y$.

3. **Unnormalized Importance Sampling.**

Consider estimating $\alpha = \mathbb{E}_\pi\phi(X)$ with $\pi(x) \propto f(x)$ via the unnormalized importance sampling estimator:

$$\tilde{\alpha}_t := \frac{\sum_{k=1}^t W_k\phi_k}{\sum_{k=1}^t W_k},$$

where $W_k := W(X_k) = f(X_k)/g(X_k), \phi_k := \phi(X_k)$, and $X_1, \ldots, X_n \sim_{\text{iid}} g(x)$. Show that the large-sample variance is

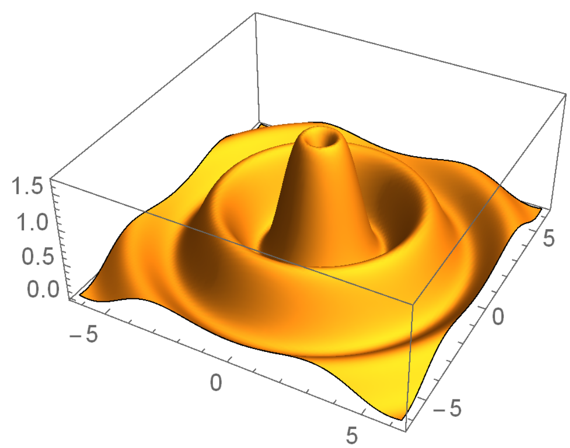
$$\mathbb{V}\text{ar}(\tilde{\alpha}_t) \simeq \frac{\mathbb{E}Z^2}{(\mathbb{E}W)^2\, t},$$

where $Z_k := W_k \times (\phi_k - \alpha)$. Hint: use the result from previous question.

4. **Numerical Experiment.** Consider the two-dimensional pdf

$$\pi(x_1, x_2) = \frac{\mathrm{e}^{-\frac{1}{4}\sqrt{x_1^2+x_2^2}}\left(\sin\left(2\sqrt{x_1^2+x_2^2}\right)+1\right)}{\alpha}, \quad \boldsymbol{x} \in [-2\pi, 2\pi]^2,$$

where α is an unknown normalization constant. The surface of this pdf (unnormalized) is depicted below. Use the importance sampling density $g(\boldsymbol{x}) = \exp(-(|x_1| + |x_2|)/4)/64$ to estimate the constant α using an estimator $\widehat{\alpha}_t$ with $t = 10^5$. Estimate the variance of $\widehat{\alpha}_t$ using the $t = 10^5$ random samples from g.



Theorem 1 (Delta Approximation)**.** *Suppose that*

$$\sqrt{t}(\boldsymbol{X}_t - \boldsymbol{\mu}) \xrightarrow{\mathrm{d}} \mathsf{N}(\mathbf{0}, \Sigma) ,$$

where $\Sigma \in \mathbb{R}^{n\times n}$*. Let* $\boldsymbol{g} : \mathbb{R}^n \mapsto \mathbb{R}^m$ *be a continuously differentiable function. Then,*

$$\sqrt{t}(\boldsymbol{g}(\boldsymbol{X}_t) - \boldsymbol{g}(\boldsymbol{\mu})) \xrightarrow{\mathrm{d}} \mathsf{N}(\mathbf{0}, \mathbf{J}\Sigma\mathbf{J}^\top) ,$$

where $\mathbf{J}$ *is the Jacobi matrix of* $\boldsymbol{g}$*,*

$$\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{x}}\right]^\top := \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix},$$

evaluated at $\boldsymbol{\mu}$*.*

Answers:

1. Trivial, to be revised in class.

2. The Jacobian matrix with $g(x, y) = x/y$ is: $[1/y, -x/y^2]$. Therefore, $\mathbf{J} = [1/\mu_Y, -\mu_X/\mu_Y^2]$ and

$$\sigma^2 := [1/\mu_Y, -\mu_X/\mu_Y^2]\Sigma \begin{bmatrix} 1/\mu_Y \\ -\mu_X/\mu_Y^2 \end{bmatrix},$$

which after simplification yields $\frac{\mathbb{V}\mathrm{ar}(X)-2\alpha\mathbb{C}\mathrm{ov}(X,Y)+\alpha^2\mathbb{V}\mathrm{ar}(Y)}{\mu_Y^2}$.

3. This is a ratio of averages estimator: $\frac{\frac{1}{t}\sum_{k=1}^t W_k\phi_k}{\frac{1}{t}\sum_{k=1}^t W_k}$ with

$$\alpha = \frac{\mathbb{E}\left[\frac{1}{t}\sum_{k=1}^t W_k\phi_k\right]}{\mathbb{E}\left[\frac{1}{t}\sum_{k=1}^t W_k\right]} = \frac{\mathbb{E}[W\phi]}{\mathbb{E}[W]}\ .$$

Hence, using the previous question with $\bar{X}_t := \frac{1}{t}\sum_{k=1}^t W_k\phi_k$ and $\bar{Y}_n := \frac{1}{t}\sum_{k=1}^t W_k$, we obtain

$$\sigma^2 = \frac{\mathbb{V}\mathrm{ar}(W\phi) - 2\alpha\mathbb{C}\mathrm{ov}(W\phi, W) + \alpha^2\mathbb{V}\mathrm{ar}(W)}{(\mathbb{E}W)^2},$$

which simplifies to $\mathbb{E}Z^2/(\mathbb{E}W)^2$.

4. Note that simulating from g is equivalent to sampling twice independently from the $\mathsf{Laplace}(0, 4)$ pdf. The $\mathsf{Laplace}(\mu, \sigma)$ density is of the form:

$$\exp\left(-\frac{|x-\mu|}{\sigma}\right)/(2\sigma)\ .$$

If $X = \sigma\mathrm{sign}(U)\ln(1-2|U|)$, where $U \sim \mathsf{U}(-1/2, 1/2)$, then $X \sim \mathsf{Laplace}(\mu, \sigma)$. Try it in R/Matlab ! Using the code below, I get $\hat{\alpha}_t \approx 52.1917$ with an estimated variance of 0.0416.

```
clear all, rand('seed',1)
f=@(X1,X2) exp(-sqrt(X1.^2 + X2.^2)/4).*...
(sin(2*sqrt(X1.^2 + X2.^2)) + 1).*(-2*pi<X1).*(X1<2*pi).*(-2*pi<X2).*(X2<2*pi);
p=@(x,s)(exp(-abs(x)/s)/(2*s));
t = 10^5; %sample size
alpha = nan(t,1);
s=4;
for i=1:t
    U=rand(1,2)-1/2;
    X = s*sign(U).*log(1-2*abs(U));
    alpha(i) = f(X(1),X(2))/(p(X(1),s)*p(X(2),s));
end
mean(alpha)
var(alpha)/t
% true answer
%q = integral2(f,-2*pi,2*pi,-2*pi,2*pi)
```