# STA 6866 Monte Carlo Statistical Methods
# Spring 2014

**Homework #3**
Assigned Monday February 3, 2014
Due Monday February 10, 2014

For the problems below, the choice of method is up to you.

1 Estimate
$$\int_{-\infty}^{\infty} \exp(-x^4)\,dx,$$
and provide an estimate of the standard error of your estimate.

**Solution** A reasonable approach is as follows. Write the integral as

$$I = \int_{-\infty}^{\infty} \frac{\exp(-x^4)}{\exp(-x^2)} \exp(-x^2)\,dx = \int_{-\infty}^{\infty} \frac{\exp(-x^4)}{\exp(-x^2)/\sqrt{\pi}} \frac{\exp(-x^2)}{\sqrt{\pi}}\,dx,$$

i.e. $I = E(h(X))$, where $X \sim \mathcal{N}(0, 1/\sqrt{2})$, and

$$h(x) = \frac{\exp(-x^4)}{\exp(-x^2)/\sqrt{\pi}}.$$

Therefore, for $X_1, \ldots, X_n \overset{\text{iid}}{\sim} \mathcal{N}(0, 1/\sqrt{2})$, we can use

$$\hat{I} = \frac{1}{n}\sum_{i=1}^{n} h(X_i),$$

with variance estimated by the sample variance of $h(X_1), \ldots, h(X_n)$.

This is implemented in R as follows.

```
set.seed(1)
ss <- 1e5
x <- rnorm(n=ss, mean=0, sd=1/sqrt(2))
h <- exp(-x^4) / (exp(-x^2)/sqrt(pi))
mean(h)                                  # gives 1.810059
sd(h)/sqrt(length(x))                    # gives 0.001692702
```

One word of caution. In principle, we need to check that the random variable $h(X)$ has a second moment with respect to the $\mathcal{N}(0, 1/\sqrt{2})$ density, or else the standard deviation estimate won't be valid. But in this particular case, we see that $h(X)$ is bounded by a constant, so there is no problem.

2 Suppose $X \sim \mathcal{N}(\theta, 1)$ and our prior on $\theta$ is the standard Cauchy distribution, with density

$$p(\theta) = \frac{1}{\pi(1+\theta^2)},$$

which is not conjugate for this situation. Estimate the posterior expectation of $\theta$ given $X = 1$, and provide an estimate of the standard error of your estimate.

**Solution** The problem can be solved either by using the rejection method or by doing importance sampling with unknown normalizing constant. Here, we discuss only the method of importance sampling with unknown normalizing constant. The posterior distribution of $\theta$ is

$$p_X(\theta) = c\frac{1}{\sqrt{2\pi}} \exp\bigl(-(X-\theta)^2/2\bigr) \, \frac{1}{\pi(1+\theta^2)}.$$

This expression suggests that we use the normal distribution with mean $X$ and standard deviation 1 (which would be the posterior if the prior was flat). The ratio of the two densities is then $(1+\theta^2)^{-1}$, except for a constant. Therefore, using the development in class (see page 4 of the notes) if we generate $\theta_1, \ldots, \theta_n \overset{\text{iid}}{\sim} N(X, 1)$, then with $h(\theta) = \theta$ and $l(\theta) = (1+\theta^2)^{-1}$ we estimate the posterior expectation via

$$\frac{\frac{1}{n}\sum_{i=1}^n h(X_i)l(X_i)}{\frac{1}{n}\sum_{i=1}^n l(X_i)}. \tag{1}$$

To estimate the variance of (1) let $V_i = h(X_i)l(X_i)$ and $W_i = h(X_i)l(X_i)$. The multivariate CLT says that

$$\frac{1}{\sqrt{n}}\Bigl(\bigl(\textstyle\sum_{i=1}^n V_i, \sum_{i=1}^n V_i\bigr) - (\mu_V, \mu_W)\Bigr) \xrightarrow{d} N(0, \Sigma)$$

where $(\mu_V, \mu_W)$ is the mean of $(V, W)$ and $\sigma$ is the covariance matrix of $(V, W)$. The estimate (1) is $g(\sum_{i=1}^n V_i, \sum_{i=1}^n V_i)$, where $g(v, w) = v/w$. Let $\nabla g(v, w) = (\partial g/\partial v, \partial g/\partial w)^\top$ be the gradient vector of $g$. The multivariate delta method says that

$$\frac{1}{\sqrt{n}}\Bigl(g\bigl(\textstyle\sum_{i=1}^n V_i, \sum_{i=1}^n V_i\bigr) - g(\mu_V, \mu_W)\Bigr) \xrightarrow{d} N\bigl(0, \nabla g(\mu_V, \mu_W)^\top \Sigma \nabla g(\mu_V, \mu_W)\bigr) \tag{2}$$

Note that $g(\mu_V, \mu_W) = \mu_V/\mu_W$ is precisely the quantity that we are trying to estimate. The variance in (2) is estimated from the $X_i$'s by plugging in sample quantities for $\mu_V$, $\mu_W$ and $\Sigma$.