Lecture 16

Error Estimates for the Monte Carlo Method

These notes address the accuracy of Monte Carlo methods in estimating probabilistic outputs. We will begin with errors when estimating the expected (i.e. mean) value, and then move on to other quantities such as the variance and probability.

16.1 Mean

Let the output of interest be labelled y, e.g. $y = T_{mh}$ in the previous turbine blade heat transfer problem. For a Monte Carlo simulation of sample size, N, label the individual values from each trial be labelled, y_i where $i = 1$ to N. In this section, we will consider the error made when estimating the expected value, also known as the mean, of y when using a sample of size N. This Monte Carlo error estimate is in fact a direct application of the field of statistics. If the distribution of y is $f(y)$, then the expected value of y is,

$$\mu_y \equiv E(y) = \int_{-\infty}^{+\infty} y f(y) dy.$$

Using the N trials, a reasonable **estimator** for μ_y would be,

$$\overline{y} \equiv \frac{1}{N} \sum_{i=1}^{N} y_i.$$

In-class Discussion 16.1 (Mean of T_{mh} from Monte Carlo) *The variability of the sample mean of T_{mh} will be demonstrated in class using various sample sizes.*

As seen in the In-class Discussion 16.1, the sample mean, $\overline{T_{mh}}$ varies from Monte Carlo simulation-to-simulation. However, as the sample size N increased, the variability of the sample means decreased. One question which might be asked is: on average how accurate is $\overline{y}$ as an estimate of μ_y? To see this, take the expectation of $\overline{y} - \mu_y$:

$$E(\overline{y} - \mu_y) \quad = \quad E(\overline{y}) - \mu_y,$$

$$
\begin{aligned}
&= E\left(\frac{1}{N}\sum_{i=1}^{N} y_i\right) - \mu_y, \\
&= \frac{1}{N}\sum_{i=1}^{N} E(y_i) - \mu_y.
\end{aligned}
$$

Since the y_i's occur from a random sampling of the inputs when using the Monte Carlo method, then $E(y_i) = \mu_y$, thus:

$$E(\overline{y} - \mu_y) = \frac{1}{N} N\mu_y - \mu_y = 0.$$

This result shows that on average, the error in using $\overline{y}$ to approximate μ_y is zero. When an estimator gives an expected error of zero, it is called an **unbiased estimator**.

To quantify the variability in $\overline{y}$, we use the variance of $\overline{y} - \mu_y$:

$$
\begin{aligned}
E\left[(\overline{y} - \mu_y)^2\right] &= E\left[\left(\left(\frac{1}{N}\sum_{i=1}^{N} y_i\right) - \mu_y\right)^2\right], \\
&= E\left[\left(\frac{1}{N}\sum_{i=1}^{N} (y_i - \mu_y)\right)^2\right], \\
&= E\left[\frac{1}{N^2}(y_1 - \mu_y + y_2 - \mu_y + \cdots)(y_1 - \mu_y + y_2 - \mu_y + \cdots)\right], \\
&= E\left[\frac{1}{N^2}\left\{(y_1 - \mu_y)^2 + (y_2 - \mu_y)^2 + \cdots + 2(y_1 - \mu_y)(y_2 - \mu_y) + \cdots\right\}\right].
\end{aligned}
$$

Because the Monte Carlo method draws independent, random samples, then the following two conditions hold,

$$E\left[(y_i - \mu_y)^2\right] = E\left[(y - \mu_y)^2\right] \equiv \sigma_y^2.$$

$$E\left[(y_i - \mu_y)(y_i - \mu_y)\right] = 0.$$

Thus, the variance of the mean estimate is,

$$E\left[(\overline{y} - \mu_y)^2\right] = \frac{\sigma_y^2}{N}.$$

Summarizing, we have found that,

$$\mu_{\overline{y}} \equiv E(\overline{y}) = \mu_y, \qquad \sigma_{\overline{y}}^2 \equiv E[(\overline{y} - \mu_y)^2] = \frac{\sigma_y^2}{N}.$$

Note, the quantity, $\sigma_{\overline{y}}$ is known as the **standard error** of the estimator. Thus, the standard error decreases with the square root of the sample size, $\sqrt{N}$. In other words, to decrease the variability in the estimate by a factor of 10 requires a factor of 100 increase in the sample size.

In-class Discussion 16.2 (Distribution of $\overline{T_{mh}}$) *In-class we will discuss how $\overline{T_{mh}}$ is distributed from Monte Carlo simulation-to-simulation.*

For large sample size N, the central limit theorem can be applied to approximate the distribution of $\overline{y}$. Specifically, the central limit theorem says for large N, the distribution of $\overline{y}$ will approach a normal distribution with mean μ_y and variation $\sigma_y/\sqrt{N}$:

$$f(\overline{y}) \rightarrow N(\mu_{\overline{y}}, \sigma_{\overline{y}}) = N(\mu_y, \sigma_y/\sqrt{N}).$$

We can now use this to make some very precise statements about the error in $\overline{y}$. Suppose, for example, that we want to have 95% confidence on the possible values for μ_y. Since the normal distribution has approximately 95% of its values within ± 2 standard deviations of the mean, then we know that,

$$P\left\{-2\frac{\sigma_y}{\sqrt{N}} \leq \bar{y} - \mu_y \leq 2\frac{\sigma_y}{\sqrt{N}}\right\} \approx 0.95.$$

If even higher confidence is wanted, then a wider range of error must be accepted. For example, a 99% confidence interval occurs at ± 3 standard deviations for a normal distribution. Thus,

$$P\left\{-3\frac{\sigma_y}{\sqrt{N}} \leq \bar{y} - \mu_y \leq 3\frac{\sigma_y}{\sqrt{N}}\right\} \approx 0.99.$$

Unfortunately, in a practical situation, we cannot actually calculate the above error estimates or confidence intervals because they depend on σ_y and we do not know σ_y. So, we typically use an estimate of σ_y. In particular, an unbiased estimate of σ_y^2 is,

$$s_y^2 \equiv \frac{1}{N-1}\sum_{i=1}^{N}(y_i - \overline{y})^2. \tag{16.1}$$

So, the usual practice is to replace σ_y by s_y in the various error estimates. Note, this does introduce additional uncertainty in the quality of the estimate and for small sample sizes this could be significant.

16.2 Other Estimators and Standard Errors

16.2.1 Probability

Often, Monte Carlo simulations are used to estimate the probability of an event occurring. For example, in the turbine blade example, we might be interested in the probability that the hot metal temperature exceeds a critical value. Generically, suppose that the event of interest is A. Then, an estimate of $P\{A\}$ is the fraction of times the event A occurs out of the total number of trials,

$$\hat{p}(A) = \frac{N_A}{N},$$

where N_A is the number of times A occurred in the Monte Carlo simulation of sample size N. $\hat{p}(A)$ is actually an unbiased estimate of $P\{A\}$. To see this, define a function $I(A_i)$

which equals 1 if event A occurred on the i-th trial, and equals zero if A did not occur. For example, if the event A is defined as $y > y_{limit}$, $I(A_i)$ would be defined as,

$$I(A_i) = I(y_i > y_{limit}) = \begin{cases} 1 & \text{if } y_i > y_{limit}, \\ 0 & \text{if } y_i \leq y_{limit}. \end{cases}$$

Using this definition, the number of times which A occurred can be written,

$$N_A = \sum_{i=1}^{N} I(A_i). \tag{16.2}$$

Finding the expectation of N_A gives,

$$\begin{aligned} E[N_A] &= E\left[\sum_{i=1}^{N} I(A_i)\right], \\ &= \sum_{i=1}^{N} E\left[I(A_i)\right]. \end{aligned}$$

Since we assume that the Monte Carlo trials are drawn at random and independently from each other, then $E[I(A_i)] = P\{A\}$. Thus,

$$E[N_A] = NP\{A\}.$$

Finally, using this result it is easy to show that,

$$E[\hat{p}(A)] = \frac{E[N_A]}{N} = P\{A\}.$$

We can also use Equation (16.2) in combination with the central limit theorem to show that $hatp(A)$ is normally distributed for large N with mean $P\{A\}$ and standard error,

$$\sigma_{\hat{p}} = \sqrt{\frac{P\{A\}(1 - P\{A\})}{N}}.$$

In-class Discussion 16.3 (Low Probability Estimation with Monte Carlo)

16.2.2 Variance

The variance of y is given the symbol, σ_y^2, and is defined as,

$$\sigma_y^2 = E[(y - \mu_y)^2]. \tag{16.3}$$

As noted in Equation (16.1), an unbiased estimator of σ_y^2 is s_y^2, that is:

$$E[s_y^2] = \sigma_y^2.$$

Note, you should try proving this result.

To quantify the uncertainty in this estimator, we would like to determine the standard error,

$$\sigma_{s_y^2} \equiv \left\{ E\left[(s_y^2 - \sigma_y^2)^2 \right] \right\}^{1/2}.$$

Unfortunately, this standard error is not known for general distributions of y. However, if y has a normal distribution, then,

$$\sigma_{s_y^2} = \frac{\sigma_y^2}{\sqrt{N/2}}.$$

Under the assumption of y being normally distributed, the distribution of s_y^2 is also related to the chi-squared distribution. Specifically, $(N-1)s_y^2/\sigma_y^2$ has a chi-square distribution with $N-1$ degrees of freedom. Note that the requirement that y be normally distributed is much more restrictive than the requirements for the mean error estimates to hold. For the mean error estimates, the standard error, $\sigma_{\overline{y}} = \sigma_y/\sqrt{N}$, is exact regardless of the distribution of y. The application of the central limit theorem which gives that $\overline{y}$ is normally distributed only requires that the number of samples is large but does constrain the distribution of y itself (beyond requiring that $f(y)$ is continuous).

16.2.3 Standard Deviation

Typically, the standard deviation of y is estimated using s_y, i.e. the square root of the variance estimator. This estimate, however, is biased,

$$E[s_y] \neq \sigma_y.$$

The standard error for this estimate is only known exactly when y is normally distributed. In that case,

$$\sigma_{s_y} \equiv \left\{ E[(s_y - \sigma_y)^2] \right\}^{1/2} = \frac{\sigma_y}{\sqrt{2N}}.$$

16.3 Bootstrapping