Simulation 3

MX

2026-01-01

Chap 7 Monte Carlo Methods in Inference

7.1 Introduction

Monte Carlo methods may refer to any method ... where simulation is used. This chapter introduces some of the Monte Carlo methods for statistical inference.

In statistical inference there is uncertainty in an estimate. To investigate this uncertainty, this chapter covers methods that use repeated sampling from a given probability model, sometimes called parametric bootstrap.

Other MC methods such as (nonparametric) bootstrap, are based on resampling from an observed sample. Resampling methods are covered in Chapters 8 and 10.

7.2 MC Methods for Estimation

Suppose $X_1, ..., X_n$ is a random sample from the distribution of X. An estimator $\hat{\theta}$ for a parameter θ is an n-variate function of the sample, $\hat{\theta} = g(X_1, ..., X_n)$.

For simplicity let $x = (x_1, ..., x_n)^T \in R^n$, and let $x^{(1)}, x^{(2)}, ...$ denote a sequence of independent random samples generated from the distribution of X. **Random variates from the sampling distribution of $\hat{\theta}$ can be generated by repeatedly drawing independent random samples $x^{(j)}$ and computing the estimate $\hat{\theta}^{(j)} = g(x_1^{(j)}, ..., x_n^{(j)})$ for each sample.**

7.2.1 MC Estimation and Standard Error

Ex 7.1. Suppose that X_1, X_2 are iid from a standard normal distribution. Estimate the mean difference $E|X_1 - X_2|$.

Here the parameter is $\theta = E(|X_1 - X_2|)$. Random variates from the sampling distribution of $\hat{\theta} = |X_1 - X_2|$ can be generated by drawing samples $x^{(j)} = (x_1^{(j)}, x_2^{(j)})$ and computing the estimate $\hat{\theta}^{(j)} = |x_1^{(j)} - x_2^{(j)}|$ and the mean of replicates, $\hat{\theta}_{MC} = \frac{1}{m}\sum_{j=1}^{m} \hat{\theta}^{(j)}$.

(Note here the estimator is $\hat{\theta} = |X_1 - X_2|$ for Monte Carlo simulation (?), not $\frac{1}{C_n^2}\sum_{1 \le i < j \le n} |X_i - X_j|$.)

The variance of $\hat{\theta}_{MC} = \frac{1}{m}\sum_{j=1}^{m} \hat{\theta}^{(j)}$ is given by $\frac{1}{m}Var(\hat{\theta}^{(j)}) = \frac{1}{m}Var(|X_1 - X_2|)$, and the Monte Carlo estimate of $Var(|X_1 - X_2|)$ is $\frac{1}{m}\sum_{j=1}^{m}(\hat{\theta}^{(j)} - \hat{\theta}_{MC})^2$

```
m = 1000
g = numeric(m)
for (i in 1:m){
  x = rnorm(2)
  g[i] = abs(x[1] - x[2])
}
theta_hat = mean(g)
var_thetahat = 1/m * mean((g - mean(g))^2)
se = sqrt(var_thetahat)
cat('MC estimate of the mean:', theta_hat, 'MC estimate of the se:', se)
```

```
## MC estimate of the mean: 1.122372 MC estimate of the se: 0.02687074
```

One can derive exact values by integration that $E|X_1 - X_2| = 2/\sqrt{\pi} = 1.128379$ and $Var(|X_1 - X_2|) = 2 - 4/\sqrt{\pi}$. The variance of the estimator is $\frac{1}{m}Var(|X_1 - X_2|) = \frac{1}{m}(2 - 4/\sqrt{\pi})$, and so the standard error is $\sqrt{\frac{1}{m}(2 - 4/\sqrt{\pi})} = 0.02695850$

7.2.2 MC Estimation of MSE

Recall that the MSE of an estimator $\hat{\theta}$ for a parameter θ is defined by $MSE(\hat{\theta}) = E[(\hat{\theta} - \theta)^2]$. If m random samples $x^{(1)}, ..., x^{(m)}$ are generated from the distribution of X, then a Monte Carlo estimate of the MSE of $\hat{\theta}$ is $\widehat{MSE} = \frac{1}{m}\sum_{j=1}^{m}(\hat{\theta}^{(j)} - \theta)^2$

Example 7.2. Obtain a Monte Carlo estimate of the MSE of the 1^{st} level trimmed mean (see the last page for definition), assuming the sampled distribution is standard normal.

Denote the estimator, trimmed sample mean, as $T := \bar{X}_{[-1]}$ and $MSE(T) = E[(T - 0)^2] = E[T^2]$. An MC estimate of $MSE(T)$ is $\widehat{MSE}(T) = \frac{1}{m}\sum_{j=1}^{m}(T^{(j)} - 0)^2$, where $T^{(j)} = \frac{1}{n-2}\sum_{i=2}^{n-1} x_{(i)}^{(j)}$ ("Random variates from the sampling distribution of $\hat{\theta}$ can be generated by repeatedly drawing independent random samples $x^{(j)}$ and computing the estimate $\hat{\theta}^{(j)} = g(x_1^{(j)}, ..., x_n^{(j)})$ for each sample.")

```r
n = 20
m = 1000
tmean = numeric(m)
for (i in 1:m){
  x = sort(rnorm(n))
  tmean[i] = sum(x[2:n-1]) / (n-2)
}
mse_mc = mean(tmean^2)
mse_mc
```

```
## [1] 0.07200666
```

7.2.3 Estimating a Confidence Level

7.3 MC Methods for Hypothesis Tests

If the test procedure is replicated a large number of times under the conditions of the null hypothesis, the observed Type I error rate should be at most (approximately) α. α is the significance level of a test, the upper bound on the probability of Type I error.

7.3.1 Empirical Type I Error Rate

The test procedure is replicated a large number of times under the null hypothesis. The empirical Type I error rate is the sample proportion of significant test statistics among the replicates.

Formally, for replicate j, generate a random sample $x_1^{(j)}, ..., x_n^{(j)}$, compute the statistic T_j, and $I_j = 1$ if H_0 is rejected and 0 otherwise. The observed Type I error rate is $\frac{1}{m}\sum_{j=1}^{m} I_j$.

Here the parameter is a probability, and the estimator is a sample proportion $\hat{p}$. The variance of $\hat{p}$ is $Var(\hat{p}) = Var(\frac{1}{m}\sum_{j=1}^{m} I_j) = \frac{1}{m}Var(I) = \frac{p(1-p)}{m}$, and so $se(\hat{p}) = \sqrt{\frac{\hat{p}(1-\hat{p})}{m}}$.

Ex 7.7. Suppose Suppose that $X_1, ..., X_{20}$ is a random sample from a $N(\mu, 100^2)$ distribution. Test $H_0 : \mu = 500$ vs. $H_1 : \mu > 500$ at $\alpha = 0.05$.

The test statistic is $T = \frac{\bar{X} - 500}{S/\sqrt{20}} \sim$ t(19) under H_0.

```
n = 20
m = 10000
mu = 500
sigma = 100
alpha = 0.05
t_crit = qt(alpha, df= n-1, lower.tail = F)

T = numeric(m)
for (j in 1:m){
  x = rnorm(n, mu, sigma)
  T[j] = (mean(x) - mu) / (sd(x)/sqrt(20))
}
p_hat = sum(T > t_crit) / m
se_phat = sqrt(p_hat * (1-p_hat) / m)
cat('observed error rate:', p_hat, 'SE of the estimator:', se_phat)
```

```
## observed error rate: 0.0519 SE of the estimator: 0.002218251
```

Close to the nominal value $\alpha = 0.05$. The proportion will vary within its standard error.

7.3.2 Power of a Test

Skipped in the main text

Trimmed mean definition: Suppose that $X_1, ..., X_n$ is a random sample and $X_{(1)}, ..., X_{(n)}$ is the corresponding ordered sample. The trimmed sample mean is computed by averaging all but the largest and smallest sample observations. More generally, the k^{th} level trimmed sample mean is defined by $\bar{X}_{[-k]} = \frac{1}{n-2k} \sum_{i=k+1}^{n-k} X_{(i)}$.