# Lecture 21: Variance Reduction Methods and Sensitivity Analysis

Today's Topics

1. Bootstrapping
2. Variance reduction methods
3. Importance sampling
4. Sensitivity analysis

# 1. Bootstrapping

- How do we get estimates of the standard errors in our estimators that don't have known distributions? e.g., in our estimate for the variance.

# 2. Variance Reduction Methods

- used to increase the accuracy of the Monte Carlo estimates that can be obtained for a given number of iterations
- "tricks" to make our MCS more "statistically efficient“
  - more accuracy for a given number of samples, or fewer samples to achieve a given level of accuracy
- variance reduction methods: importance sampling, antithetic sampling, control variates, stratified sampling

# 3. Importance Sampling

- a general technique for estimating properties of one distribution while only having samples generated from another (different) distribution

# 3.1 Importance sampling general idea

→ A technique for estimating properties of one distribution while only having samples generated from another (different) distribution

Consider random variable $X$, pdf $f_X(x)$

$$\mu_X \equiv E_X[X] = \int x \, f_X(x) \, dx$$

↑ denotes expectation under $f_X$

Choose a random variable $Z \geq 0$

$$\text{s.t.} \quad E_X[Z] = 1 = \int \underbrace{z \, f_X(x)}_{f_Z(x)} \, dx$$

# 3.1 Importance sampling general idea

Then $\mu_x = E_x[X] = \int x f_x(x)\, dx$

$$= \int \frac{x}{z} z \, f_x(x)\, dx$$

$$= \int \frac{x}{z} f_z(x)\, dx = E_z\left[\frac{x}{z}\right]$$

# 3.1 Importance sampling general idea

We have

$$E_X[X] = E_Z\left[\frac{X}{Z}\right]$$

sample $x$ under $f_X(x)$

estimator variance is $\frac{var_X[X]}{N}$

sample $\frac{x}{z}$ under $f_Z(x)$

estimator variance is $\frac{var_Z\left[\frac{X}{Z}\right]}{N}$

# 3.1 Importance sampling general idea

We have

$$E_X[X] = E_Z\left[\frac{X}{Z}\right]$$

sample $x$ under $f_X(x)$ — sample $\frac{X}{Z}$ under $f_Z(x)$

estimator variance is $\dfrac{\mathrm{var}_X[X]}{N}$ — estimator variance is $\dfrac{\mathrm{var}_Z\left[\frac{X}{Z}\right]}{N}$

- Some values of $X$ in our MCS have more impact on the parameter being estimated (here $\mu_X$) than others
- If "important" values are emphasized by sampling more fequently, then estimator variance can be reduced
- Key is to choose an appropriate "biasing distribution"
  → good choice means we can decrease $N$ for same accuracy

# 3.2 Importance sampling for probability estimation

We saw $P\{A\}$ is estimated via MCS with $\hat{p}(A)$.

Consider $A$ is the event $y > y_{limit}$ (eg. prob. of failure)

Define indicator function $I(A_i) = \begin{cases} 1, & y_i > y_{limit} \\ 0, & y_i \leq y_{limit} \end{cases}$

↑ sample $i$

Then $\hat{p}(A) = \frac{1}{N} \sum_{i=1}^{N} I(A_i)$

and $E[\hat{p}(A)] = P\{A\}$ (unbiased)

$$\mathrm{var}[\hat{p}(A)] = \frac{P\{A\}(1 - P\{A\})}{N}$$

# 3.2 Importance sampling for probability estimation

Introduce pdf $w(x)$ (alternative pdf for $x$)

→ called "biasing density"

→ choose $w(x)$ so that event $A$ occurs more frequently

Then $P\{A\} = E_x[I(A)] = \int I(A) f(x)\, dx$

expectation under $X$

$$= \int I(A) \frac{f(x)}{w(x)} w(x)\, dx$$

$$= E_w\left[ I(A) \frac{f(x)}{w(x)} \right]$$

draw from pdf $w(x)$

weight by $\frac{f(x)}{w(x)}$ to counter

# 3.2 Importance sampling for probability estimation

Then our MC IS estimator for $P\{A\}$ is

$$\hat{p}_{IS}(A) = \frac{1}{N} \sum_{i=1}^{N} I(A_i) \frac{f(x_i)}{w(x_i)}$$

where the $x_i$ are drawn from $w(x)$

<u>Summary</u>:

- define $w(x)$
- draw samples ~~the~~ of $x$ from $w(x)$
  (idea → more samples in region of interest)
- estimate $P\{A\}$, but we need to weight the samples by $\frac{f(x_i)}{w(x_i)}$ to account for the fact that we drew from $w(x)$ not $f(x)$

Then:

$$E\left[\hat{p}_{IS}(A)\right] = P\{A\} \quad \text{(unbiased)}$$

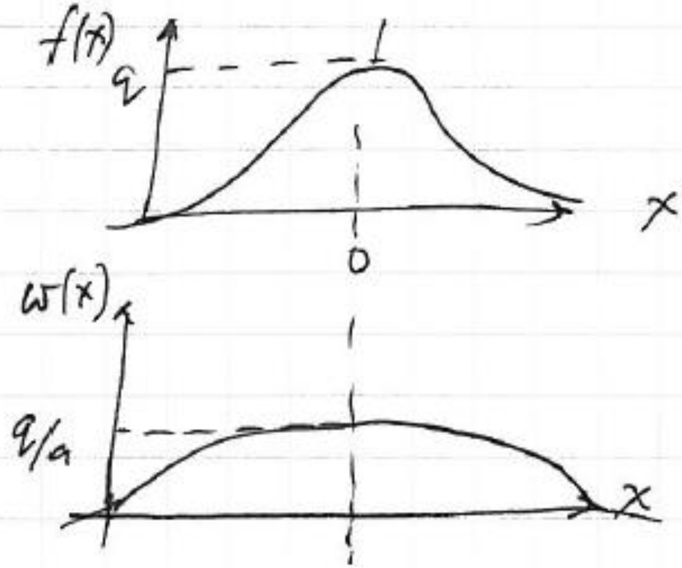
$$\mathrm{var}_w\left[\hat{p}_{IS}(A)\right] = \frac{1}{N}\left\{E\left[I(A)\frac{f(x)}{w(x)}\right] - \left(P\{A\}\right)^2\right\}$$

# 3.3 How to pick the biasing distribution

Simple approach: Scaling to shift probability mass into the event region

$$w(x) = \frac{1}{a} f\left(\frac{x}{a}\right)$$

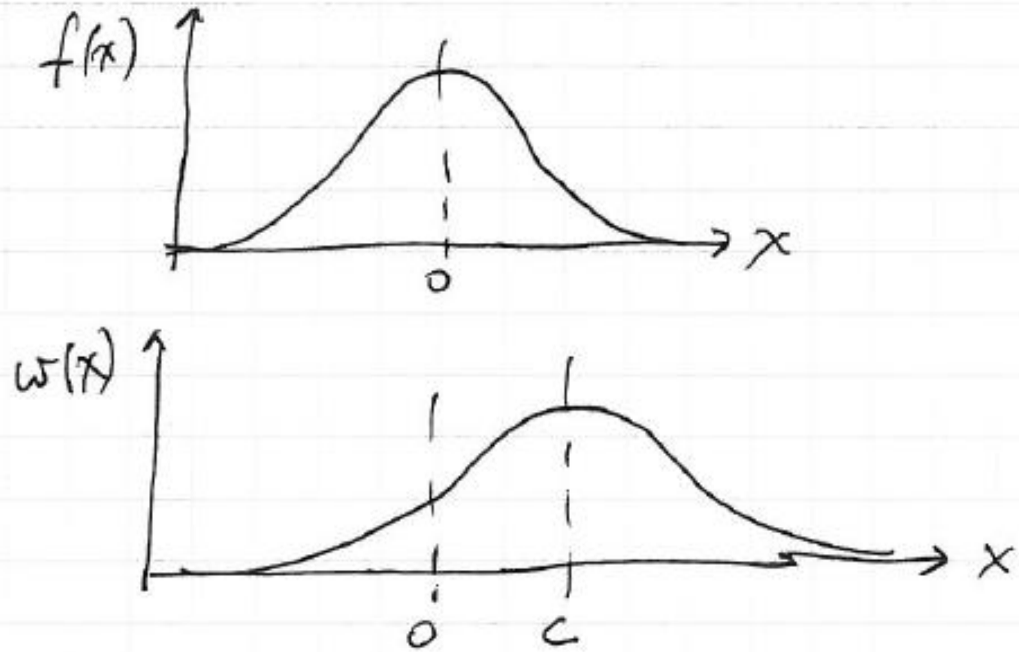
$a > 1$




fatter tails

# 3.3 How to pick the biasing distribution

translation: $w(x) = f(x-c)$, $c > 0$ (for $P(x > x_{crit})$)



# 4. Sensitivity Analysis

- How do we use our MCS results to understand which uncertain inputs are contributing the most to output variability?
  - Important to understand where we should focus our uncertainty reduction efforts (improving manufacturing tolerances, improving models, installing sensors, etc.)
    → factor prioritization
  - Important to understand where there may be uncertainties but they are not important
    → factor fixing

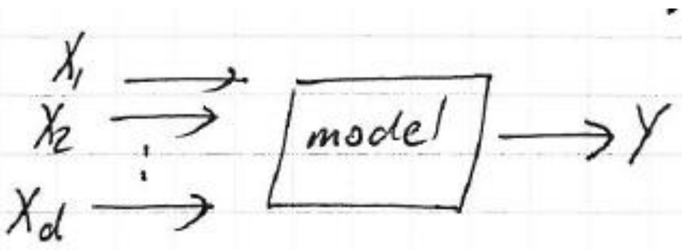
# 4a. Vary-all-but-one (VABO) MCS

1. Run a MCS with all inputs varying
2. Fix input k to a deterministic value. Rerun the MCS with all inputs except k varying.
3. Compare statistics of the output (e.g., variance of Run #1 to variance of Run #2).

Questions:

- at what value should we fix factor k?
- would the results be different if we fixed factor k to a different deterministic value?
- what about possible interactions among the inputs?

# 4b. Global Sensitivity Analysis




d random inputs

1 random output.

Main effect sensitivity index for input $i$ : $S_i = \dfrac{var(Y) - E[var(Y|X_i)]}{Var(Y)}$

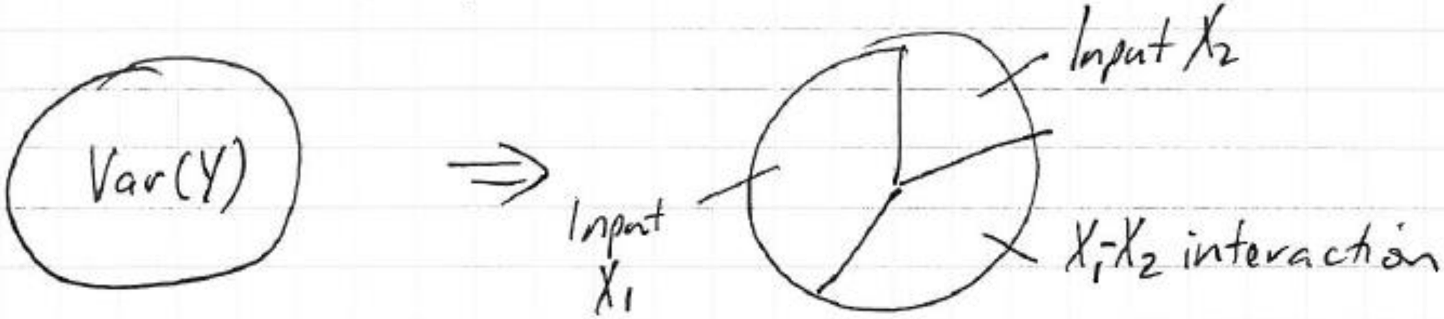
→ expected, relative reduction in output variance if the true value of $X_i$ is learned (expectation over what that true value might be)

→ measure of the effect of varying $X_i$ alone, averaged over variations in other inputs

# 4b. Global Sensitivity Analysis

Also can compute higher-order interaction indices

$S_{ij}$, $S_{ijk}$, ... etc.




$$S_1 + S_2 + S_{12} = 1$$

Generally $\sum_{i=1}^{d} S_i + \sum_{i<j}^{d} S_{ij} + \cdots + S_{12\ldots d} = 1$

Total effect sensitivity index for input $i$ : $S_{T_i} = \dfrac{var(Y) - E[var(Y \mid X_{\sim i})]}{var(Y)}$

→ measures contribution to var(Y) of $X_i$ including its main effect and all the interaction effects