ANOTHER IMPORTANCE SAMPLER:

Compute $\theta = \int_0^1 e^{x^2} dx$

now we use
$g(x) = 1 + x^2 + \frac{x^4}{2}$

The normalization is

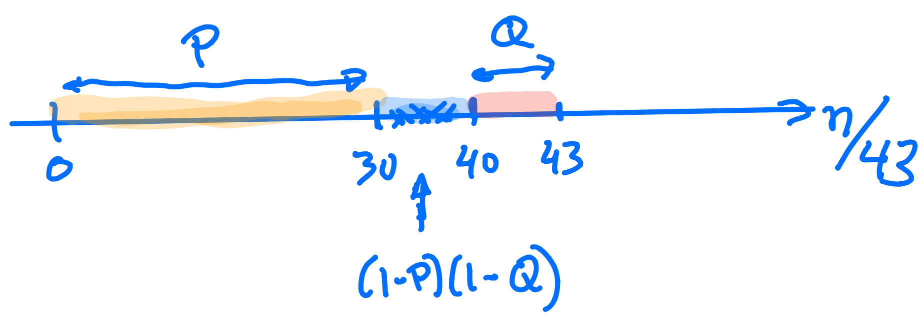
$Z = \int_0^1 g(x)\,dx = \frac{43}{30}$. The CDF is then

$$G(x) = \frac{30}{43}\int_0^x g(t)\,dt = \frac{30}{43}x + \frac{10}{43}x^3 + \frac{3}{43}x^5$$

$X_i \sim G(x)$

Draw $U \sim \mathcal{U}(0,1)$ $\quad V \sim \mathcal{U}(0,1)$

let $P < \frac{30}{43}\mathcal{U}(0,1)$ $\quad Q > \frac{1}{43}\mathcal{U}(40,43)$



$$X_i = PU + (1-P)(1-Q)U^{1/3} + QU^{1/5}$$

$$\hat{\theta} = \frac{43}{30} \int_0^1 \frac{e^{x^2}}{1+x^2+x^4/2} \, \frac{30(1+x^2+x^4/2)}{43} \, dx$$

$$\approx \frac{43}{40N} \sum_{i=1}^{N} \frac{e^{X_i^2}}{1+X_i^2+X_i^2/2}$$

GENERATION OF SAMPLES NUMERICALLY:

Commonly no direct F^{-1}. We resort to numerics: goal is to find the x such that

$$(*) \quad F(x) = u, \qquad u \in \mathcal{U}(0,1)$$

$$f(x) = F'(x) \text{ is the pdf}$$

Could use Newton Root finding. It is fast, but can fail. Other methods: Regula Falsi, Secant, fixed point.

Newton. Suppose we have a guess x^0. Set up the iteration

$$(\ddagger) \qquad F(x^\ell + \delta x^\ell) = u \quad , \ell = 0, 1, 2 \ldots$$

Let $x = \lim_{\ell \to \infty} x^\ell$ is the root.

If $|\delta x^\ell| << 1$ then

$$F(x^\ell + \delta x^\ell) \approx F(x^\ell) + \frac{dF}{dx}(x^\ell)\, \delta x^\ell$$

OK for $|\delta x^\ell| << 1$. Since $\frac{dF}{dx} = f(x)$

$$F(x^\ell) + f(x^\ell)\, \delta x^\ell = u$$. We know x^ℓ

$$f(x^\ell)\, \delta x^\ell = u - F(x^\ell).$$

Solve for δx^ℓ.

then update $x^{\ell+1} = x^\ell + \delta x^\ell$, iterate.

Require $f'(x)$ continuous.

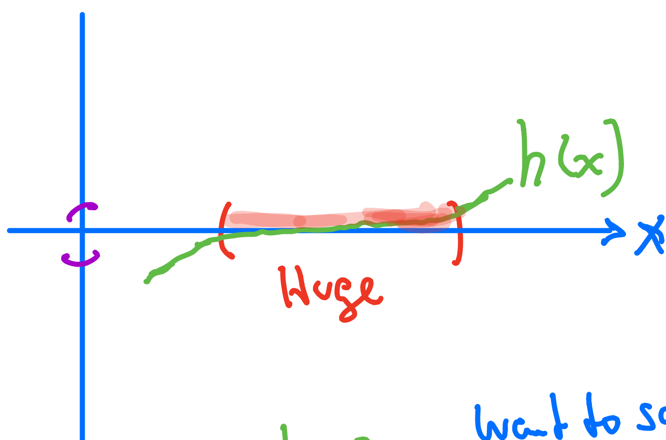
Convergence rate: $\delta x^\ell = x^{\ell+1} - x^\ell$

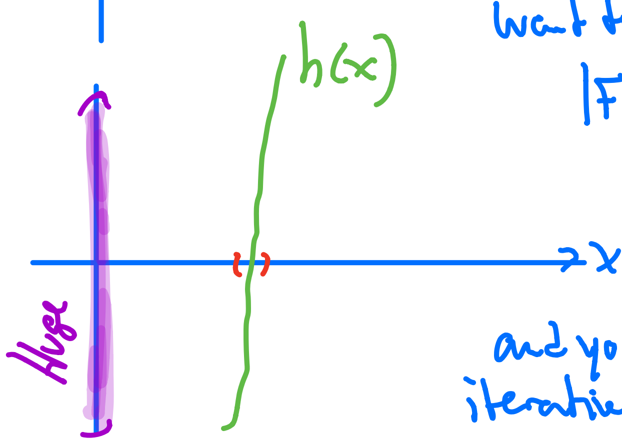
$$|x^{l+1} - x^{l}| = C\,|x^{l} - x|^{2}$$

quadratic convergence.

Problems when $f \sim 0$ ... a large step occurs. This leads to severe numerical problems. Actually, there are 2 problems.

Let $h(x) = F(x) - u = 0$





want to satisfy BOTH

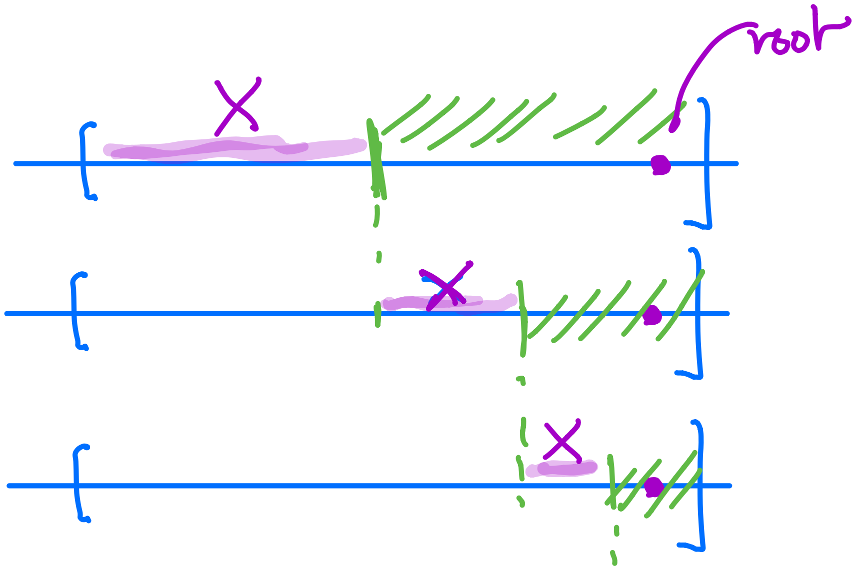
$|F(x) - u| < \varepsilon$

and

$|\delta x| < \delta$

and you add a maximum iteration count as well

IF Newton Fails: Bisection is slowest, but robust.

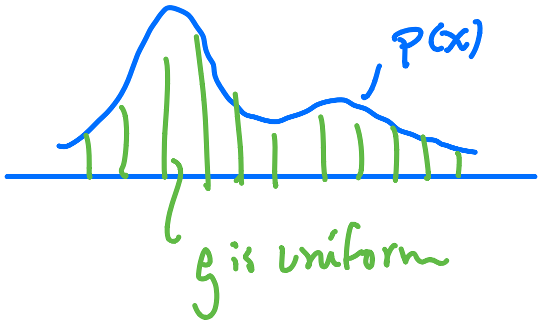


and so on.

$$\text{So} \quad \phi = \mathbb{E}(f) = \int_{\Omega} f(x)\, p(x)\, dx$$

$$\hat{\phi} = \frac{1}{N} \sum_{i=1}^{N} \frac{f(x^i)\, p(x^i)}{g(x^i)} \qquad x_i \sim G(x)$$

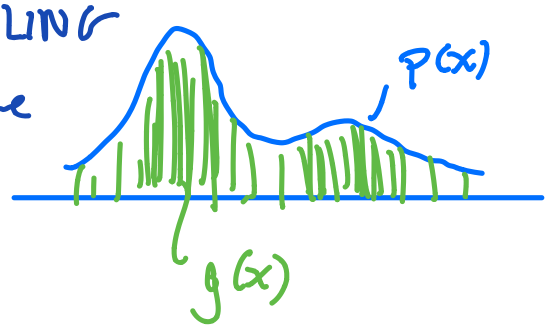
want g to be as close to p so that ratio looks like a uniform distribution: equal weights $w_i = p(x^i)/g(x^i)$



REGULAR MC SAMPLING

IMPORTANCE SAMPLING

put more samples where p is largest.



Also, want g(x) to have same or larger support than p(x): you dont want p(X)/g(X) to give nonsensical answers.

SELF NORMALIZED IMPORTANCE SAMPLING

Want to compute $\phi = \int f(x) p_0(x)\, dx$, $p_0(x)$ is not necessarily a pdf, i.e. $p_0(x) \geq 0$ but $Z_p = \int p_0(x)\, dx$. So $p(x) \equiv \frac{p_0(x)}{Z_p}$

$$\phi = \int \frac{f(x)\, p_0(x)\, g(x)}{g(x)}\, dx$$

$$= Z_p \int \frac{f(x)\, p(x)\, g(x)}{g(x)}\, dx$$

$$= \frac{Z_p \int f(x) \frac{p(x)\, g(x)}{g(x)}\, dx}{Z_p \int \frac{\cancel{p}\, g}{g}\, dx} = \frac{\int f(x) \frac{p_0}{g} g\, dx}{\int \frac{p_0}{g} g\, dx}$$

$$= \frac{\int f(x)\, w(x)\, g(x)\, dx}{\int w(x)\, g(x)\, dx} = \frac{\mathbb{E}_g(fw)}{\mathbb{E}_g(w)}$$

$$\boxed{w = \frac{p_0}{g}}$$

$$\hat{\phi} = \frac{\sum_{i=1}^{N} f(x^i)\, w(x^i)}{\sum_{i=1}^{N} w(x^i)}$$

$$x^i \sim G(x)$$

If all the weight is concentrated into a few of the $w_i = w(x^i)$ then the calculation will have severe errors.

Consider the following toy problem: (forget $f(x)$): let Z_i be IID $i = 1, 2, \dots, n$.
They all have variance σ^2

An estimator of the mean

$$\hat{\mu} = \frac{\sum_{i=1}^{N} w_i Z_i}{\sum_{i=1}^{N} w_i}$$

$$\mathrm{Var}(\hat{\mu}) = \sigma^2 \frac{\left(\sum w_i\right)^2}{\sum w_i^2}$$

$$\therefore \mathrm{var}\, \hat{\mu} = \frac{\sigma^2}{n_e}$$

$$n_e = \frac{\left(\sum w_i\right)^2}{\sum w_i^2}$$

n_e is then a figure of merit of

the effective samples.

Trivial example: suppose $w_i = 1$ for k of the weights $\Rightarrow$ only k samples are effective $n_e = k$. That is, it's like averaging with only k instead of n samples.