A General Acceptance-Rejection Technique for Probability Distributions

i. A general purpose variate construction technique (von Neumann, 1951) when $f(x)$ is known but other **variate generation techniques** (construction, inversion of $F(x)$, approximation by $H(x)$, and numerical solutions to $u = F(x)$) fail outright, are difficult, or computationally prohibitive.

ii. This technique provides a reasonably efficient random variate for the $Gamma(a,1)$ distribution, from which variates for $Erlang(n,b)$ and $Chisquare(n)$ can be derived which **are not** $O(n)$.

A General Acceptance-Rejection Technique for Probability Distributions

CASE STUDY: A recent paper presented research which required the use of **reservoir sampling** of HUGE data collections. As j becomes large, many, many "pulls" are required from the pRNG to service the $Bernoulli(n/(j+1))$ test.

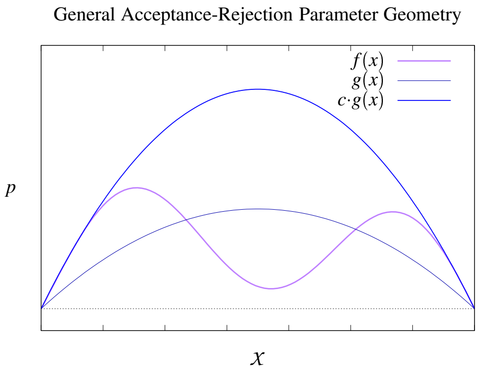
This was inefficient for pRNGs "big enough" to support their study sampling requirements,[1] so they used this technique to generate samples from a $Geometric(p)$ **like**[2] distribution... Instead of $10,000$ pRNG calls to find the next population element to keep, they made only 8–10.

$\mathcal{S} \leftarrow$ `first` n `elements of` $\mathcal{P}$
$j \leftarrow n$
while `(` `elements exist in` $\mathcal{P}$ `)` **do**
    $e \leftarrow$ `next element from` $\mathcal{P}$
    if `(` $Bernoulli\left(\frac{n}{j+1}\right)$ `is True` `)` **then** `(`
        $i \leftarrow Equilikely(0, n-1)$
        $\mathcal{S}[i] \leftarrow e$
    `)`
    `increment` j
`)`

[1] We'll talk about this and other pRNG details in lecture shortly...

[2] Recall $Geometric(p)$ is the number of failures before the first success of $Bernoulli(p)$ trials. In reservoir sampling p changes with every *Bernoulli* trial, so the distribution is not as straight forward as *Geometric*.

General Acceptance-Rejection Theorem — Implementation



Given $G^{-1}(u)$ the IDF for $g(x)$, $f(x)$, and $c > 0$ such that $f(x) \leq c \cdot g(x)$ over $\mathcal{X}$:

repeat (
  $u \leftarrow Random()$
  $x \leftarrow G^{-1}(u)$
  $v \leftarrow Random()$
) **until** ($c \cdot g(x) \cdot v \leq f(x)$)

x is a sample from $\mathcal{X}$ proportionally drawn with $f(x)$

General Acceptance-Rejection Theorem — Summary

- An accept-reject algorithm for arbitrary probability distributions with ellusive or inefficient $F^{-1}(x)$ implementations.
- Requires (at least) two $Random()$ calls and a $G^{-1}(u)$ evaluation per iteration

Given $G^{-1}(u)$ the IDF for $g(x)$, $f(x)$, and $c > 0$ such that $f(x) \leq c \cdot g(x)$ over $\mathcal{X}$:

repeat (
    $u \leftarrow Random()$
    $x \leftarrow G^{-1}(u)$
    $v \leftarrow Random()$
) **until** ($c \cdot g(x) \cdot v \leq f(x)$)

x is a sample from $\mathcal{X}$
  proportionally drawn with $f(x)$

General Acceptance-Rejection Theorem — Summary

- An accept-reject algorithm for arbitrary probability distributions with ellusive or inefficient $F^{-1}(x)$ implementations.
- Requires (at least) two $Random()$ calls and a $G^{-1}(u)$ evaluation per iteration
- What do we look for in a $g(x)$ and G^{-1}? Why not simply use

$$c \cdot g(x) = \max_{\mathcal{X}} f(x)$$

in which case $g(x)$ is the uniform distribution over $\mathcal{X}$.

Given $G^{-1}(u)$ the IDF for $g(x)$, $f(x)$, and $c > 0$ such that $f(x) \leq c \cdot g(x)$ over $\mathcal{X}$:

repeat (
$u \leftarrow Random()$
$x \leftarrow G^{-1}(u)$
$v \leftarrow Random()$
) **until** ($c \cdot g(x) \cdot v \leq f(x)$)

x is a sample from $\mathcal{X}$ proportionally drawn with $f(x)$

General Acceptance-Rejection Theorem — Summary

- An accept-reject algorithm for arbitrary probability distributions with ellusive or inefficient $F^{-1}(x)$ implementations.
- Requires (at least) two $Random()$ calls and a $G^{-1}(u)$ evaluation per iteration
- What do we look for in a $g(x)$ and G^{-1}? Why not simply use

$$c \cdot g(x) = \max_{\mathcal{X}} f(x)$$

in which case $g(x)$ is the uniform distribution over $\mathcal{X}$.

Given $G^{-1}(u)$ the IDF for $g(x)$, $f(x)$, and $c > 0$ such that $f(x) \leq c \cdot g(x)$ over $\mathcal{X}$:

repeat (
  $u \leftarrow Random()$
  $x \leftarrow G^{-1}(u)$
  $v \leftarrow Random()$
) **until** ($c \cdot g(x) \cdot v \leq f(x)$)

x is a sample from $\mathcal{X}$ proportionally drawn with $f(x)$

First, and importantly, if $f(x)$ has an infinite tail, $g(x) = 0$ over $\mathcal{X}$ and we can't possibly find a satisfying $c > 0$.

General Acceptance-Rejection Theorem — Summary

- An accept-reject algorithm for arbitrary probability distributions with ellusive or inefficient $F^{-1}(x)$ implementations.
- Requires (at least) two $Random()$ calls and a $G^{-1}(u)$ evaluation per iteration
- What do we look for in a $g(x)$ and G^{-1}? Why not simply use

  $$c \cdot g(x) = \max_{\mathcal{X}} f(x)$$

  in which case $g(x)$ is the uniform distribution over $\mathcal{X}$.

Given $G^{-1}(u)$ the IDF for $g(x)$,
$f(x)$, and $c > 0$ such that
$f(x) \leq c \cdot g(x)$ over $\mathcal{X}$:

repeat (
  $u \leftarrow Random()$
  $x \leftarrow G^{-1}(u)$
  $v \leftarrow Random()$
) **until** ($c \cdot g(x) \cdot v \leq f(x)$)

x is a sample from $\mathcal{X}$
proportionally drawn with $f(x)$

More pragmatically, we want c to be as small as possible, in other words we want $g(x)$ to closely match $f(x)$ — since larger cs means more reject loops through the failed predicate

$$c \cdot g(x) \cdot v \leq f(x) \Rightarrow v \leq \frac{f(x)}{c \cdot g(x)}$$