شبیه سازی تصادفی

تولید اعداد تصادفی

۳ مهر ۱۳۹۹

تولید اعداد تصادفی

- توزیع یکنواخت
- نظریه اعداد
- آزمایش اعداد تصادفی
- توصیه های مولدهای عدد تصادفی

خلاصه

- ما درباره شبه اعداد تصادفی صحبت می کنیم
- تعداد زیادی RNG یا مولدهای اعداد تصادفی وجود دارد
- ... با کیفیت متفاوت
- موارد خاص خود را اجرا نکنید، مگر برای سرگرمی یا تحقیقات پروژه.
- RNG داخلی باید قبل از استفاده مورد بررسی قرار گیرد
- ... حداقل در محیط های توسعه هدف کلی
- محیط های محاسبات علمی معمولا از پیشرفته ترین سطح RNG برخوردار است که می تواند قابل اعتماد باشد.
- اگر شرایط به اندازه کافی شدید باشد، هر RNG شکست خواهد خورد.

- نیاز مشهود به اعداد تصادفی
- جدول ها
- ژنراتورهای فیزیکی.دستگاه های قرعه کشی

تعریف

- توزیع یکنواخت [۰ ,۱].
- تصادفی بودن (استقلال)
- یک مشکل اساسی این است که کامپیوترها در اعداد تصادفی حقیقی کار نمیکنند:دنباله ای از متغیر تصادفی مستقل U_i ، به طور یکنواخت روی [۰،۱] توزیع شده
- دنباله ای از اعداد مستقل و دارای توزیع یکنواخت در بازه [۰،۱]

تعریف یک مولد اعداد تصادفی

تعریف

یک مولد اعداد تصادفی،یک الگوریتم رایانه ای است که دنباله ای از اعداد حقیقی یا صحیح را است. به صورت تصادفی تولید می کند که به طور یکنواخت روی $[۰ ; ۱]$ یا $\{۰, ..., N - ۱\}$ توزیع شده اند و از لحاظ آماری مستقل اند.

روش حذفی میان مربعی برای تولید اعداد تصادفی

تعریف

در این روش برای تولید اعداد با k رقم اعشار، در آغاز یک عدد k رقمی را به عنوان هسته برمی گزینیم و آن را با x_0 نشان می دهیم.سپس آن را به توان دو می رسانیم.عدد حاصل حداکثر دارای ۲k رقم است؛در صورتی که تعداد ارقام عدد کمتر از ۲k باشد،سمت چپ عدد تعدادی صفر قرار می دهیم تا ۲k رقمی شود.سپس اگر k زوج باشد، $\frac{k}{2}$ رقم را از دو طرف عدد و اگر فرد باشد، $\frac{(k+1)}{2}$ رقم را از سمت راست و $\frac{(k-1)}{2}$ رقم را از سمت چپ آن حذف می کنیم.این عدد k رقمی را با x_1 نمایش داده ئ آن را مانند x_0 به توان دو می رسانیم و روند بالا را دنبال می کنیم.

مثال

برای k=۴ و $x_0 = ۵۴۹۷$ با روش میان مربعی ، سه مقدار اول دنباله به صورت زیر خواهند بود:

$$x_0^2 = ۵۴۹۷^2 = ۳۰۲۱۷۰۰۹ \Rightarrow x_1 = ۲۱۷۰ \Rightarrow u_1 = ۰/۲۱۷۰$$
$$x_1^2 = ۲۱۷۰^2 = ۰۴۷۰۸۹۰۰ \Rightarrow x_2 = ۷۰۸۹ \Rightarrow u_2 = ۰/۷۰۸۹$$
$$x_2^2 = ۷۰۸۹^2 = ۵۰۲۵۳۹۲۱ \Rightarrow x_3 = ۲۵۳۹ \Rightarrow u_3 = ۰/۲۵۳۹$$

- دنباله فیبوناتچی: $x_۱ = ۱ \quad x_۰ = ۰ \quad i \geq ۲ \qquad x_i = x_{i-۱} + x_{i-۲}$
- مولد فیبوناتچی: همچنین روش همنهشتی جمعی نامیده می شود.
$x_i = \mod(x_{i-۱} + x_{i-۲}, M) \quad U_i = \frac{x_i}{M}$ که $x = \mod(y, M)$ مدول
بعد از تقسیم می باشد یعنی $y - nM$ که $n = \lfloor y/M \rfloor$ و هم چنین داریم:
$x_i \in [۰, M - ۱]$
ماکزیمم طول دوره $M^۲ - ۱$ است که برای $M = ۲, ۳$ حاصل می شود.

مولد $U_i \in [۰, ۱] \quad U_i = \mathrm{mod}(aU_{i-۱}, ۱)$ نشان دهنده این است که a بزرگ است.آخرین رقم ها خفظ می شوند.به صورت زیر قابل پیاده سازی است: (x_i یک عدد صحیح است)

$$x_i = \mathrm{mod}(ax_{i-۱}, M) \quad U_i = \frac{x_i}{M}$$

مثال ها $a = ۲۳$ و $M = ۱۰^۸ + ۱$ هستند.

نتیجه گیری میانی

- حالت اولیه کل دنباله را تعیین می کند.
- تعداد چرخه های مختلف
- طول هر چرخه

اگر x_i بتواند مقادیر N را به دست آورد،حداکثر طول یک چرخه، N است.

خصوصیات یک مولد تصادفی

- طول چرخه
- تصادفی بودن
- سرعت
- قابل تجدید
- قابل حمل

مولد هم نهشتی خطی (روش هم نهشتی خطی برای تولید اعداد تصادفی)

مولد هم نهشتی خطی (LCG) به صورت زیر تعریف می شود:

$$x_i = \mathrm{mod}(ax_{i-۱} + c, M) \quad U_i = \frac{x_i}{M}$$

برای ضریب a ، انتقال c و مدول M .
ما a ، c و $x_۰$ را طوری در نظر می گیریم که x_i در $(۰, ۱, ..., M - ۱)$ نهفته باشد و تصادفی بنظر برسد.

مثال

$M=۱۶$ و $a=۵$ و $c=۱$

با $x_۰ = ۳$:

۳ ۱۰ ۵ ۴ ۷ ۱۴ ۹ ۸ ۱۱ ۲ ۱۳ ۱۲ ۱۵ ۶ ۱ ۰

قضیه

مولد همنهشتی خطی (LCG) دارای طول تام است اگر و فقط اگر

الف: M و c نسبت به یکدیگر اول باشند.

ب: برای هر عامل اول p از M ، $\gcd(a, p) = ۱$

پ: اگر ۴ عاملی از M باشد، آنگاه $\gcd(a, ۴) = ۱$. توجه داشته باشید که اگر M اول باشد، دوره کامل تنها وقتی حاصل می شود که $a=۱$ باشد.

مثال

به عنوان مثال XOR بین چندین ژنراتور

- دوره را افزایش می دهد
- تصادفی بودن را بهبود می بخشد
- اما به خوبی درک نشده است
- استفاده از LCG به طور گسترده توصیه می شود

- یک ثبت تغییر شکل بازخورد خطی بزرگ
- ۱۹۹۳۷ بیت حافظه استفاده می کند.
- حداکثر دوره را دارد.یعنی $۲^{۱۹۹۳۷۷} - ۱$
- دارای توزیع مناسب است
- همچنین توزیع مشترک ۶۲۳ بعدی
- احتمالا بهترین PRNG تاکنون برای شبیه سازی تصادفی (نه برای رمزنگاری)

RNG ها در محیط های رایج

- R
- Python
- S-plus
- Matlab

تعریف

یک دنباله از اعداد شبه تصادفی U_i ، یک دنباله قطعی از اعداد در بازه $[۰،۱]$ هستند که دارای همان خاصیت های آماری مربوطه به عنوان دنباله ای از اعداد تصادفی می باشند.

- نوع توزیع
- تصادفی بودن (استقلال،سفیدی)