Marcelo Q. A. Oliveira, Tamer S. G. Cavalcante,
Heitor S. Ramos, Osvaldo A. Rosso, Alejandro C. Frery
Laboratório de Computação Cienífica e Análise Numérica
Universidade Federal de Alagoas (LaCCAN-UFAL)

Avaliação de Geradores de Números Pseudoaleatórios Através de Técnicas da Teoria da Informação

ERAD-NE 2015

Agenda

Introdução

Motivação

- Criptografia

Motivação

- Criptografia
- Amostragem

Motivação

- Criptografia
- Amostragem
- Aplicações gráficas

Motivação

- Criptografia
- Amostragem
- Aplicações gráficas
- Games

Motivação

- Criptografia
- Amostragem
- Aplicações gráficas
- Games
- **Simulação**

Agenda

Geradores Reais

Baseados em algum fenômeno natural aleatório

- Ruído atmosférico capturado por um rádio [random.org 1998, (random.org)].

Geradores Reais

Baseados em algum fenômeno natural aleatório

- Ruído atmosférico capturado por um rádio [random.org 1998, (random.org)].
- Tempo entre emissão de partículas durante o decaimento radioativo [Walker 1998, HOTBITS]

Geradores Reais

Baseados em algum fenômeno natural aleatório

- Ruído atmosférico capturado por um rádio [random.org 1998, (random.org)].
- Tempo entre emissão de partículas durante o decaimento radioativo [Walker 1998, HOTBITS]
- Ruído térmico oriundo de semicondutores em um circuito (Intel Ivy Bridge) [Hamburg et al. 2012,] [Goodin 2013,]

Geradores Reais

Baseados em algum fenômeno natural aleatório

- Ruído atmosférico capturado por um rádio [random.org 1998, (random.org)].
- Tempo entre emissão de partículas durante o decaimento radioativo [Walker 1998, HOTBITS]
- Ruído térmico oriundo de semicondutores em um circuito (Intel Ivy Bridge) [Hamburg et al. 2012,] [Goodin 2013,]
- Monitoramento de ruído em ambientes [vanheusden.com 2012].

Geradores Reais

Baseados em algum fenômeno natural aleatório

- Ruído atmosférico capturado por um rádio [random.org 1998, (random.org)].
- Tempo entre emissão de partículas durante o decaimento radioativo [Walker 1998, HOTBITS]
- Ruído térmico oriundo de semicondutores em um circuito (Intel Ivy Bridge) [Hamburg et al. 2012,] [Goodin 2013,]
- Monitoramento de ruído em ambientes [vanheusden.com 2012].

Desvantagens

- Necessitam de Hardware específico

Geradores Reais

Baseados em algum fenômeno natural aleatório

- Ruído atmosférico capturado por um rádio [random.org 1998, (random.org)].
- Tempo entre emissão de partículas durante o decaimento radioativo [Walker 1998, HOTBITS]
- Ruído térmico oriundo de semicondutores em um circuito (Intel Ivy Bridge) [Hamburg et al. 2012,] [Goodin 2013,]
- Monitoramento de ruído em ambientes [vanheusden.com 2012].

Desvantagens

- Necessitam de Hardware específico
- Não reprodutíveis

Geradores de Números Pseudoaleatórios:

Geradores de Números Pseudoaleatórios - PRNGs:

- Algorítmicos (determinísticos)

Geradores de Números Pseudoaleatórios:

Geradores de Números Pseudoaleatórios - PRNGs:

- Algorítmicos (determinísticos)
- Produzem sequências que se comportam como as produzidas por geradores reais a partir de sementes conhecidas [L'Ecuyer 2007]

Geradores de Números Pseudoaleatórios:

Geradores de Números Pseudoaleatórios - PRNGs:

- Algorítmicos (determinísticos)
- Produzem sequências que se comportam como as produzidas por geradores reais a partir de sementes conhecidas [L'Ecuyer 2007]
- Mais convenientes por não necessitar de hardware específico

Geradores de Números Pseudoaleatórios:

Geradores de Números Pseudoaleatórios - PRNGs:

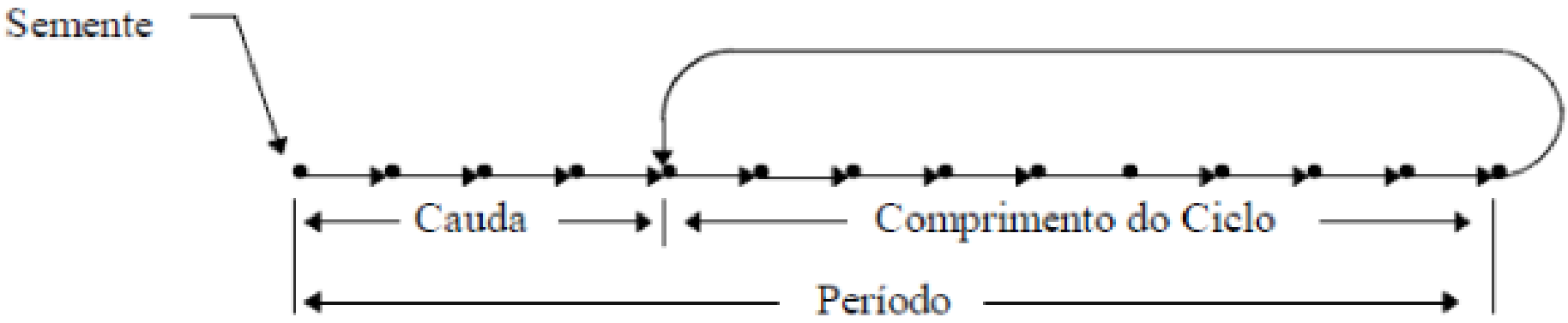
- Algorítmicos (determinísticos)
- Produzem sequências que se comportam como as produzidas por geradores reais a partir de sementes conhecidas [L'Ecuyer 2007]
- Mais convenientes por não necessitar de hardware específico
- Possibilitam a reprodutibilidade

Geradores de Números Pseudoaleatórios:

Geradores de Números Pseudoaleatórios - PRNGs:

- Algorítmicos (determinísticos)
- Produzem sequências que se comportam como as produzidas por geradores reais a partir de sementes conhecidas [L'Ecuyer 2007]
- Mais convenientes por não necessitar de hardware específico
- Possibilitam a reprodutibilidade

Período de um gerador



Geradores de Números Pseudoaleatórios PRNGs

Propriedades desejáveis em PRNGs:

- O período de repetição seja suficientemente grande.

Geradores de Números Pseudoaleatórios PRNGs

Propriedades desejáveis em PRNGs:

- O período de repetição seja suficientemente grande.
- A geração de números deve ser rápida.

Geradores de Números Pseudoaleatórios PRNGs

Propriedades desejáveis em PRNGs:

- O período de repetição seja suficientemente grande.
- A geração de números deve ser rápida.
  - Poupar recursos computacionais para as aplicações em si.

Geradores de Números Pseudoaleatórios PRNGs

Propriedades desejáveis em PRNGs:

- O período de repetição seja suficientemente grande.
- A geração de números deve ser rápida.
  - Poupar recursos computacionais para as aplicações em si.
- Os números gerados devem seguir uma distribuição uniforme.

Geradores de Números Pseudoaleatórios PRNGs

Propriedades desejáveis em PRNGs:

- O período de repetição seja suficientemente grande.
- A geração de números deve ser rápida.
  - Poupar recursos computacionais para as aplicações em si.
- Os números gerados devem seguir uma distribuição uniforme.
  - Devem ter a mesma probabilidade de ocorrência.

Geradores de Números Pseudoaleatórios PRNGs

Propriedades desejáveis em PRNGs:

- O período de repetição seja suficientemente grande.
- A geração de números deve ser rápida.
  - Poupar recursos computacionais para as aplicações em si.
- Os números gerados devem seguir uma distribuição uniforme.
  - Devem ter a mesma probabilidade de ocorrência.
- Os números devem ser estatisticamente independentes entre si.

Geradores de Números Pseudoaleatórios PRNGs

Propriedades desejáveis em PRNGs:

- O período de repetição seja suficientemente grande.
- A geração de números deve ser rápida.
  - Poupar recursos computacionais para as aplicações em si.
- Os números gerados devem seguir uma distribuição uniforme.
  - Devem ter a mesma probabilidade de ocorrência.
- Os números devem ser estatisticamente independentes entre si.
  - O valor de um número na sequência não deve afetar o valor do próximo.

Geradores de Números Pseudoaleatórios PRNGs

Propriedades desejáveis em PRNGs:

- O período de repetição seja suficientemente grande.
- A geração de números deve ser rápida.
  - Poupar recursos computacionais para as aplicações em si.
- Os números gerados devem seguir uma distribuição uniforme.
  - Devem ter a mesma probabilidade de ocorrência.
- Os números devem ser estatisticamente independentes entre si.
  - O valor de um número na sequência não deve afetar o valor do próximo.

Método Congruencial Linear (LCG)

Sejam os números uniformes inteiros $U1, U2, U3, \ldots$ entre 0 e m - 1, em que m representa um grande número inteiro. Podemos gerar estes números utilizando o método congruencial por meio da relação recursiva:

Método Congruencial Linear (LCG)

Sejam os números uniformes inteiros $U1, U2, U3, \ldots$ entre 0 e m - 1, em que m representa um grande número inteiro. Podemos gerar estes números utilizando o método congruencial por meio da relação recursiva:

LCG

$$U_i + 1 = (aU_i + c) mod\ m$$

Método Congruencial Linear (LCG)

Sejam os números uniformes inteiros $U1, U2, U3, \ldots$ entre 0 e m - 1, em que m representa um grande número inteiro. Podemos gerar estes números utilizando o método congruencial por meio da relação recursiva:

LCG

$$U_i + 1 = (aU_i + c) mod\ m$$

Onde:

- m é chamado de módulo;
- a e c, inteiros positivos denominados multiplicador e incremento respectivamente;
- mod é um operador que retorna o resto da divisão de $a_U i + c$ por m;

Mersenne Twister - MT

Mersenne Twister - Matsumoto [Matsumoto and Nishimura 1998]

- Entrega inteiros de 32 bits.

Mersenne Twister - MT

Mersenne Twister - Matsumoto [Matsumoto and Nishimura 1998]

- Entrega inteiros de 32 bits.
- Período de 2^{19937}-1.

Mersenne Twister - MT

Mersenne Twister - Matsumoto [Matsumoto and Nishimura 1998]

- Entrega inteiros de 32 bits.
- Período de 2^{19937}-1.
- Passa na maioria dos testes conhecidos.

Agenda

Suítes

- Diehard (FORTRAN) [Marsaglia 1995]

Suítes

- Diehard (FORTRAN) [Marsaglia 1995]
- NIST [NIST 1999]

Suítes

- Diehard (FORTRAN) [Marsaglia 1995]
- NIST [NIST 1999]
- Dieharder (ANSI C) [Brown et al. 2004]

Suítes

- Diehard (FORTRAN) [Marsaglia 1995]
- NIST [NIST 1999]
- Dieharder (ANSI C) [Brown et al. 2004]
- TEST U01 (biblioteca em ANSI C) [L'Ecuyer and Simard 2007]

Agenda

Plano Entropia x Complexidade

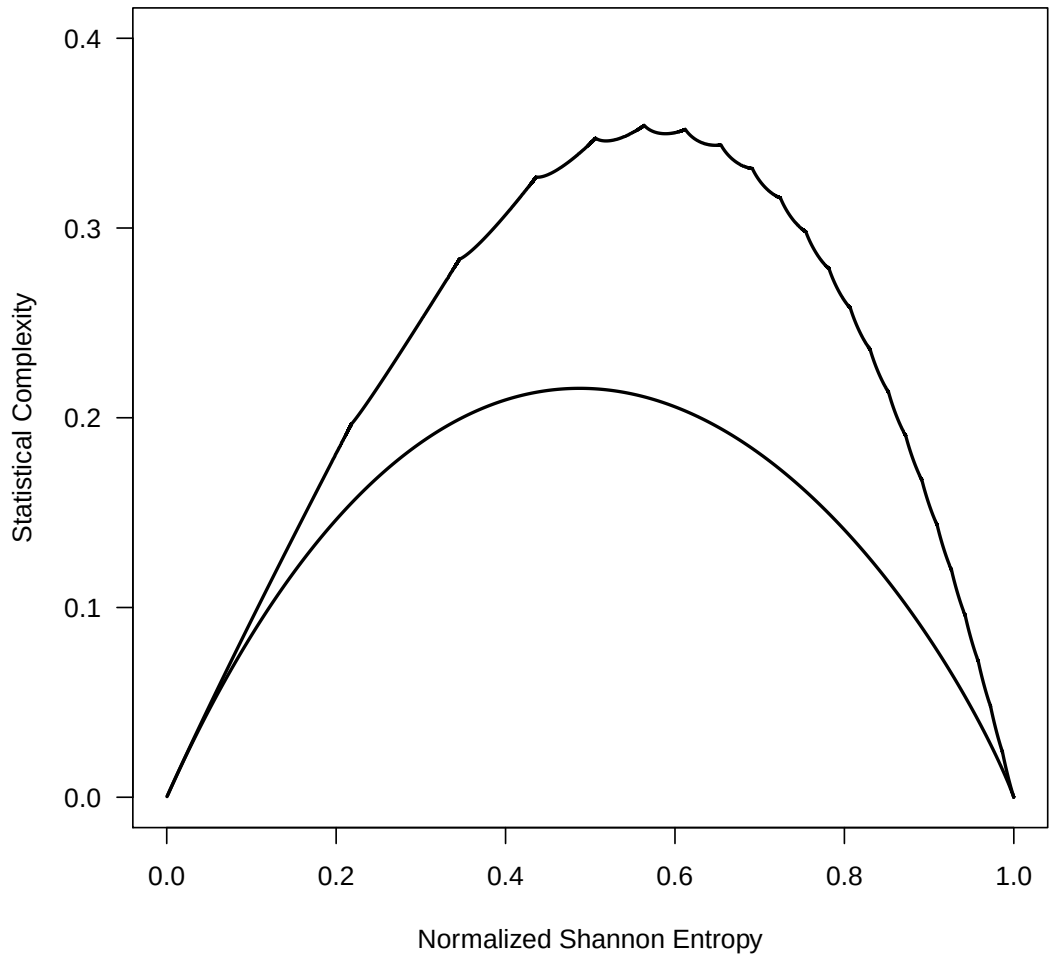


Figura 1 : Plano HC - Entropia x Complexidade Estatística

Agenda

Teste Proposto

- Teste de hipóteses não paramétrico para medir a qualidade da sequência gerada por um PRNG através da posição do ponto observado no plano (HC).

Teste Proposto

- Teste de hipóteses não paramétrico para medir a qualidade da sequência gerada por um PRNG através da posição do ponto observado no plano (HC).
- Com o objetivo de ter uma referência foram utilizados dados oriundos de um gerador real.

Teste Proposto

- Teste de hipóteses não paramétrico para medir a qualidade da sequência gerada por um PRNG através da posição do ponto observado no plano (HC).
- Com o objetivo de ter uma referência foram utilizados dados oriundos de um gerador real.
- Os dados foram fornecidos pelo grupo de Processamento de Informação Quântica do Instituto de Tecnologia Max Plank, num arquivo binário de aproximadamente 200Mb obtido segundo o processo descrito em [Gabriel et al. 2010].

Teste Proposto

- Tais dados foram mapeados como uma sequência de 108 números aleatórios no intervalo $(0, 1)$, e então particionados em 10^5 sequências de 10^3 elementos cada uma.

Teste Proposto

- Tais dados foram mapeados como uma sequência de 108 números aleatórios no intervalo $(0, 1)$, e então particionados em 10^5 sequências de 10^3 elementos cada uma.
- Posteriormente, foram calculados os valores da entropia e da complexidade estatística para cada uma das subsequências, resultando em 10^5 pontos no plano (H,C).

Teste Proposto

- Como apontado por [Larrondo et al. 2013, Larrondo] , uma sequência aleatória ideal produziria o valor $(1, 0)$ no plano HC.

Teste Proposto

- Como apontado por [Larrondo et al. 2013, Larrondo] , uma sequência aleatória ideal produziria o valor $(1, 0)$ no plano HC.
- Elaboramos um teste de hipóteses não paramétrico para medir a qualidade da sequência gerada por um PRNG qualquer através da posição do ponto observado no plano (H,C).

Teste Proposto

- Como apontado por [Larrondo et al. 2013, Larrondo] , uma sequência aleatória ideal produziria o valor $(1, 0)$ no plano HC.
- Elaboramos um teste de hipóteses não paramétrico para medir a qualidade da sequência gerada por um PRNG qualquer através da posição do ponto observado no plano (H,C).
- A medida é feita comparando o ponto com aqueles obtidos das sequências de mesmo tamanho produzidas pelo RNG descrito em [Gabriel et al. 2010].

Teste Proposto

- Como apontado por [Larrondo et al. 2013, Larrondo] , uma sequência aleatória ideal produziria o valor $(1, 0)$ no plano HC.
- Elaboramos um teste de hipóteses não paramétrico para medir a qualidade da sequência gerada por um PRNG qualquer através da posição do ponto observado no plano (H,C).
- A medida é feita comparando o ponto com aqueles obtidos das sequências de mesmo tamanho produzidas pelo RNG descrito em [Gabriel et al. 2010].
- Testamos duas sequências de tamamnho 10^3 produzidas pelos geradores Mersenne Twister (MT) e Congruencial Linear (LCG).

Agenda

Resultados

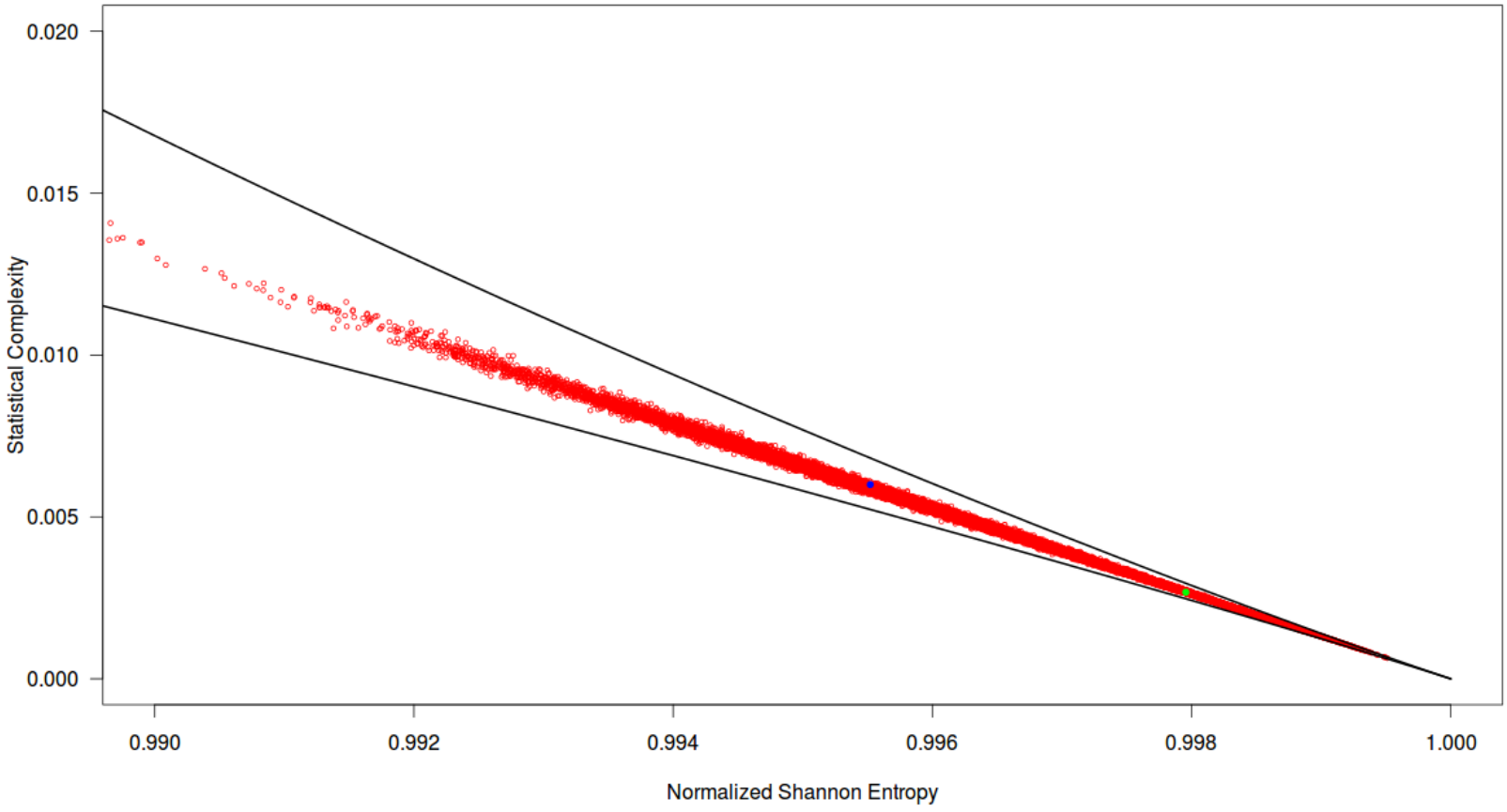


Figura 2 : 10^3 pontos no Plano HC

Resultados

P-valor

- p-valor LCG: 0.135

Resultados

P-valor

- p-valor LCG: 0.135
- p-valor MT: 0.850

Resultados

P-valor

- p-valor LCG: 0.135
- p-valor MT: 0.850

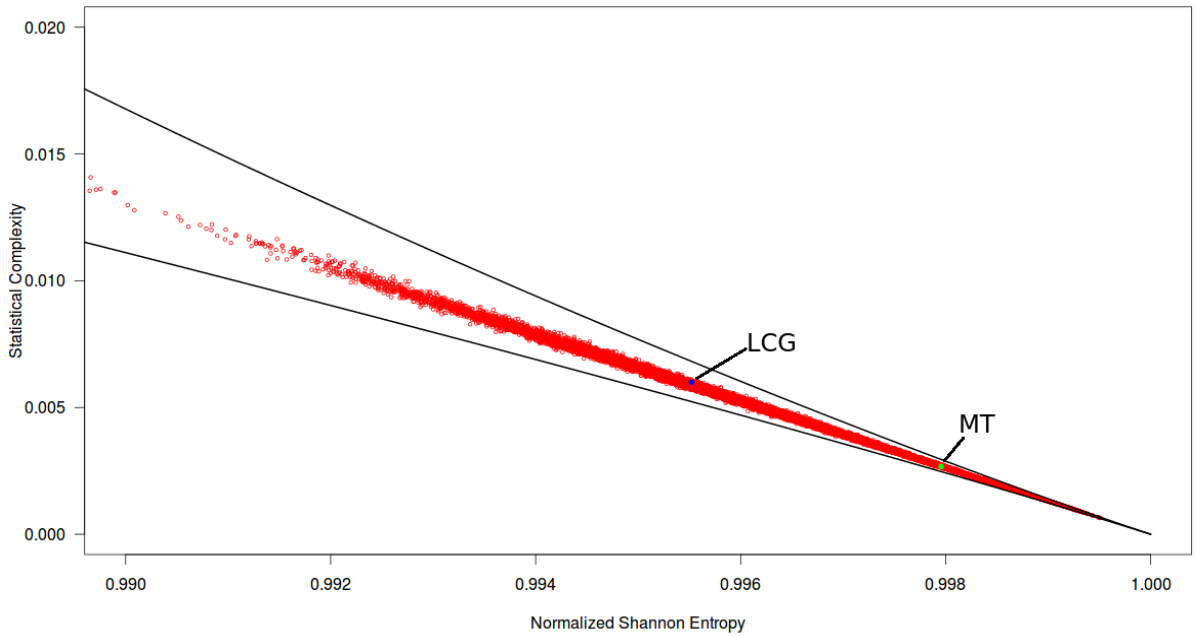


Figura 3 : Pontos LCG e MT

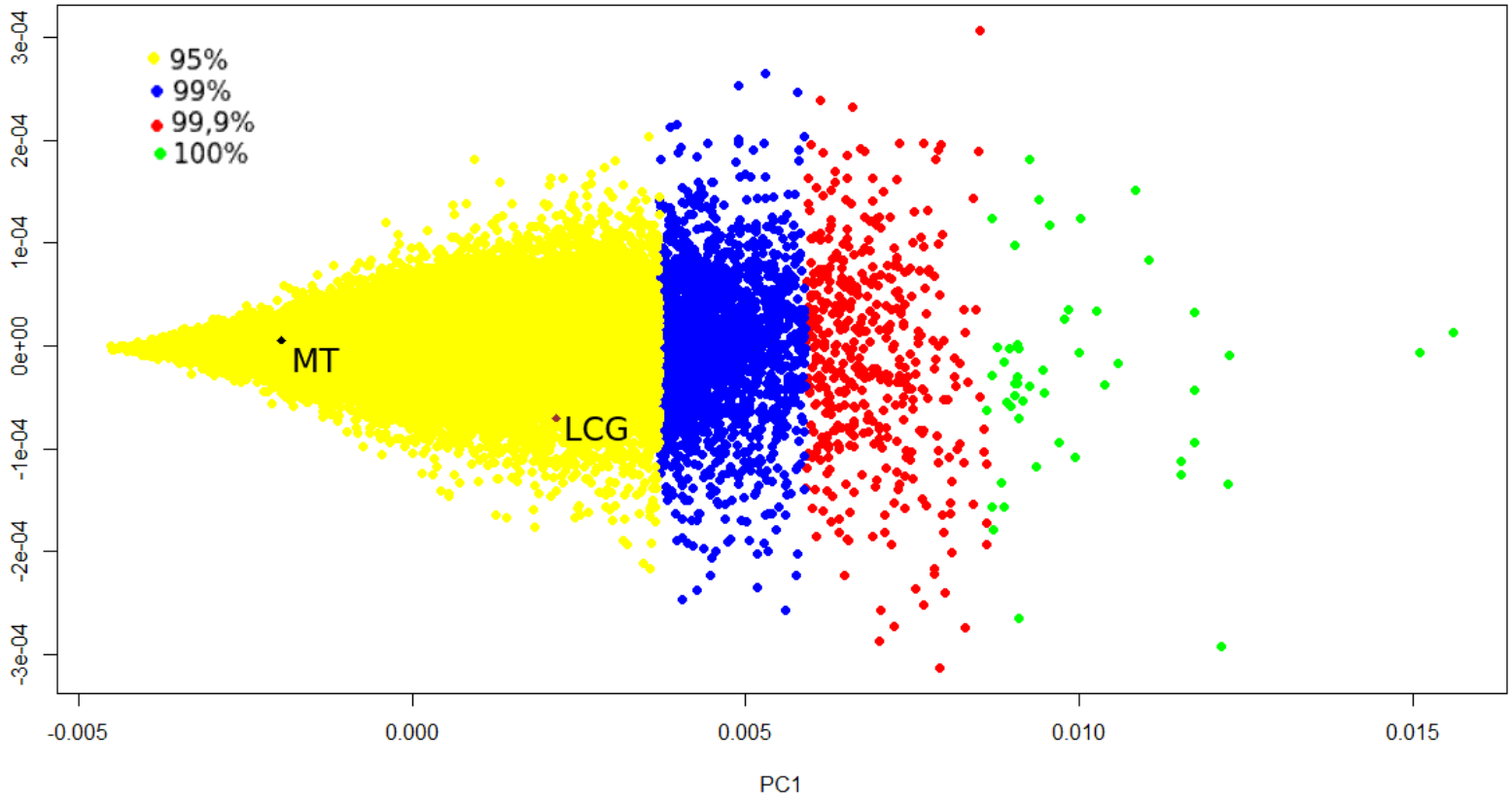


Figura 4 : Dados submetidos a PCA

Obrigado!


Brown, R. G., Eddelbuettel, D., and Bauer, D. (2004).
Dieharder: A random number test suite.
http://www.phy.duke.edu/~rgb/General/dieharder.php.
Acessado em 11/2014.


Gabriel, C., Wittmann, C., Sych, D., Dong, R., Mauerer, W., Andersen, U. L., Marquardt, C., and Leuchs, G. (2010).
A generator for unique quantum random numbers based on vacuum states.
Nature Photonics, 4(10):711–715.


Goodin, D. (2013).
Researchers can slip an undetectable trojan into intel's ivy bridge cpus.
http://arstechnica.com/security/2013/09/researchers-can-slip-an-undetectable-trojan-into-intels-ivy

Acessado em 11/2014.


Hamburg, M., Kocher, P., and Marson, M. E. (2012).

Analysis of intel's ivy bridge digital random number generator.
http://www.rambus.com/wp-content/uploads/2015/08/Intel_TRNG_Report_20120312.pdf.
Acessado em 11/2014.


Larrondo, H. A., De Micco, L., Gonzalez, C. M., Plastino, A., and Rosso, O. A. (2013).
Statistical Complexity of Chaotic Pseudorandom Number Generators | BenthamScience.
Concepts and Recent Advances in Generalized Information Measures and Statistics, pages 283–308.


L'Ecuyer, P. (2007).
Random Number Generation, pages 93–137.
John Wiley & Sons, Inc.


L'Ecuyer, P. and Simard, R. (2007).
TestU01: A C Library for Empirical Testing of Random Number Generators.
ACM Transactions on Mathematical Software, 33(4):Article 22.


Marsaglia, G. (1995).
Diehard.
http://stat.fsu.edu/pub/diehard/.
Acessado em 11/2014.


Matsumoto, M. and Nishimura, T. (1998).
Mersenne twister: A 623-dimensionally equidistributed uniform pseudo-random number generator.
ACM Trans. Model. Comput. Simul., 8(1):3–30.


NIST (1999).
Nist statistical test suite.
http://www.itl.nist.gov/div893/staff/soto/jshome.html.
Acessado em 11/2014.


random.org (1998).
True random number service.
https://www.random.org/.
Acessado em 11/2014.

vanheusden.com (2012).
audio entropy daemon.
https://www.vanheusden.com/aed/.
Acessado em 11/2014.

Walker, J. (1998).
Genuine random numbers, generated by radioactive decay.
https://www.fourmilab.ch/hotbits/.
Acessado em 11/2014.