ΤΕΧΝΙΚΕΣ ΠΡΟΣΟΜΟΙΩΣΗΣ

Επικ. Καθ. Στέλιος Ζήμερας
Τμήμα Μαθηματικών
Κατεύθυνση Στατιστικής και Αναλογιστικών – Χρηματοοικονομικών Μαθηματικών

2017

Γεννήτριες Τυχαίων Αριθμών

- Παραγωγή ακολουθίας τυχαίων αριθμών
- Μια ακολουθία από αριθμούς που παράγονται από κάποια πηγή θεωρείται τυχαία όταν ικανοποιούνται οι εξής δύο βασικές ιδιότητες:

1. Όλοι οι αριθμοί της ακολουθίας είναι ομοιόμορφα κατανεμημένοι στο πεδίο ορισμού τους.
2. Οι παραγόμενοι αριθμοί είναι στατιστικά ανεξάρτητοι.

Γεννήτριες Τυχαίων Αριθμών

- Μία γεννήτρια ψευδοτυχαίων αριθμών θα πρέπει να ικανοποιεί τις παρακάτω ιδιότητες:

1. Οι αριθμοί που παράγει πρέπει να είναι όσο το δυνατόν πιο ανεξάρτητοι μεταξύ τους. Στην ιδανική περίπτωση δεν πρέπει να υπάρχει αυτοσυσχέτιση μεταξύ τους. Δηλαδή για κάποιον που δεν γνωρίζει τον αλγόριθμο της επαναληπτικής διαδικασίας οι αριθμοί που γεννιόνται να είναι τυχαίοι.

2. Θα πρέπει να είναι γρήγορη και να μην απαιτεί μεγάλη υπολογιστική μνήμη.

3. Οι αριθμοί που παράγει θα πρέπει να έχουν μεγάλο κύκλο παραγωγής (περίοδος).

4. Θα πρέπει να έχει την δυνατότητα να αναπαράγει την ίδια ακολουθία αριθμών, ώστε να επαναλαμβάνεται η προσομοίωση του μοντέλου, για διαφορετικά εναλλακτικά σενάρια, στις ίδιες συνθήκες.

Γεννήτριες Τυχαίων Αριθμών

- **Μέθοδος Μέσων Τετραγώνων**

Μια από τις πρώτες επαναληπτικές διαδικασίες που χρησιμοποιήθηκε για την παραγωγή ψευδοτυχαίων αριθμών είναι η μέθοδος του μέσων τετραγώνων.

Κάθε νέος αριθμός της ακολουθίας των ψευδοτυχαίων παράγεται από τα ρ μεσαία ψηφία του τετραγώνου ενός ρ-ψήφιου αριθμού.

Γεννήτριες Τυχαίων Αριθμών

Η μέθοδος ακολουθεί τα εξής βήματα:

1. Επιλέγουμε ένα ρ-ψήφιο αριθμό.
2. Το αριθμό αυτό τον υψώνουμε στο τετράγωνο και προσθέτουμε μηδενικά στα αριστερά του αριθμού, αν χρειάζονται για να γίνει 2ρ-ψήφιος αριθμός.
3. Τα ρ μεσαία ψηφία του αποτελούν τον επόμενο ψευδοτυχαίο αριθμό.
4. Επαναλαμβάνουμε τα βήματα 2 και 3 για κάθε νέο ψευδοτυχαίο αριθμό που δημιουργούμε.

Γεννήτριες Τυχαίων Αριθμών

- Μειονέκτημα της μεθόδου είναι ότι η ακολουθία των ψευδοτυχαίων που δημιουργείται επαναλαμβάνει τον εαυτό της με μικρή περίοδο, δηλαδή έχουμε σύντομα επανάληψη ψευδοτυχαίου αριθμού.
- Ένα άλλο μειονέκτημα της μεθόδου είναι η εμφάνιση του αριθμού 0 σε κάποια φάση της διαδικασίας, κάτι που όμως αντιμετωπίζεται με προγραμματισμένη αντικατάστασή του.

Γεννήτριες Τυχαίων Αριθμών

i	x_i	x_i^2
1	21	0441
2	44	1936
3	93	8649
4	64	4096
5	09	0081
6	08	0064
7	06	0036
8	03	0009
9	00	0000
10	00	0000

Γεννήτριες Τυχαίων Αριθμών

- **Μέθοδος Μέσων Γινομένων**

Η μέθοδος αυτή μοιάζει με τη μέθοδο των μέσων τετραγώνων με τη μόνη διαφορά ότι κάθε νέος ψευδοτυχαίος αριθμός προκύπτει από τα μεσαία ρ

ψηφία του γινομένου των δύο προηγούμενων ρ-ψήφιων αριθμών της ακολουθίας των ψευδοτυχαίων.

Γεννήτριες Τυχαίων Αριθμών

Τα βήματα της μεθόδου είναι:

1. Αρχικά επιλέγουμε δύο ρ-ψήφιους αριθμούς.
2. Τους αριθμούς αυτούς τους πολλαπλασιάζουμε και προσθέτουμε στα αριστερά του γινομένου μηδενικά, αν χρειάζονται για να γίνει 2ρ-ψήφιος το γινόμενο.
3. Τα ρ μεσαία ψηφία του γινομένου αποτελούν τον επόμενο ψευδοτυχαίο αριθμό.
4. Πολλαπλασιάζουμε τον νέο ψευδοτυχαίο αριθμό με τον προηγούμενο αριθμό της ακολουθίας των ψευδοτυχαίων και επαναλαμβάνουμε το βήμα 3.

Γεννήτριες Τυχαίων Αριθμών

i	x_i	$x_i \cdot x_{i-1}$
1	21	
2	20	0420
3	42	0840
4	84	3528
5	52	4368
6	36	1872
7	87	8132
8	13	κ.λ.π.

Γεννήτριες Τυχαίων Αριθμών

- **Congruential Μέθοδοι (Γραμμικές γεννήτριες υπολοίπων)**

Υπάρχουν τρεις τύποι congruential μεθόδων:

1. Ο πολλαπλασιαστικός τύπος:

$$x_{n+1} = (\alpha \cdot x_n) \bmod m$$

2. Ο μικτός τύπος:

$$x_{n+1} = (\alpha \cdot x_n + c) \bmod m$$

3. Ο προσθετικός (αθροιστικός) τύπος:

$$x_{n+1} = (x_n + x_{n-1}) \bmod m$$

όπου α, c και m είναι σταθερές και x_0 είναι η αρχική τιμή (seed) που χρησιμοποιείται για τη δημιουργία της ακολουθίας των ψευδοτυχαίων.

Γεννήτριες Τυχαίων Αριθμών

- οι ψευδοτυχαίοι μπορούν να πάρουν τιμές μεταξύ 0 και m-1, δηλαδή 0 < x_i < m-1
- οι congruential μέθοδοι μπορούν να δώσουν ψευδοτυχαίους αριθμούς μεταξύ 0 και 1, δηλαδή

  0 < r_i < 1, εάν θέσουμε

$$r_i = \frac{x_i}{m}$$

Γεννήτριες Τυχαίων Αριθμών

- **πολλαπλασιαστικός τύπος**

$$x_{n+1} = (\alpha \cdot x_n) \bmod m$$

Τα x_0 και m πρέπει να πρώτοι αριθμοί μεταξύ τους.

1. Για να έχουμε καλύτερα αποτελέσματα θα πρέπει το m να είναι δύναμη του 2, δηλαδή $m = 2^b$
2. Η μέγιστη περίοδος ψευδοτυχαίων που παράγονται με τον πολλαπλασιαστικό τύπο, όταν $m = 2^b$

και ισχύουν οι συνθήκες:

ι) $\alpha \bmod 8 = \pm 3$ ή 5 και

ιι) x_0 και m πρώτοι αριθμοί μεταξύ τους,

είναι ίση με 2^{b-2}

Γεννήτριες Τυχαίων Αριθμών

Παράδειγμα:

Έστω $m = 16 = 2^4$, $\alpha = 5$ ($5 \bmod 8 = 5$) και $x_0 = 3$, τότε

i	x_i	$5x_i$	$5x_i \bmod 16$
1	3	15	15
2	15	75	11
3	11	55	7
4	7	35	3
5	3		

η περίοδος της παραγόμενης ακολουθίας των ψευδοτυχαίων είναι $4 = 2^{4-2} = 2^2$

75/16 = 4 υπόλοιπο 11

Γεννήτριες Τυχαίων Αριθμών

ΒΗΜΑΤΑ

- Ξεκινάμε με μία αρχική τιμή $x_0 \in N$ η οποία καλείται seed
- υπολογίζεται η νέα (ψευδοτυχαία) τιμή

$$x_1 = (ax_0) \bmod m$$

για κάποιους προεπιλεγμένους φυσικούς αριθμούς a,m

- υπολογίζουμε το

$$x_2 = (ax_1) \bmod m$$

$$x_3 = (ax_2) \bmod m$$

κ.ο.κ.

Γεννήτριες Τυχαίων Αριθμών

- κάθε x_i ανήκει στο {0,1,...,m−1} και ο αριθμός

$$\frac{x_i}{m} \in \left\{0, \frac{1}{m}, \ldots, \frac{m-1}{m}\right\} \subseteq [0,1)$$

μπορεί να θεωρηθεί ως ένα ψευδοτυχαίος αριθμός μεταξύ του 0 και του 1

$$U_i = \frac{x_i}{m} \sim \left(0,1\right)$$

RAND

Γεννήτριες Τυχαίων Αριθμών

Αλγόριθμος

B1. Θέτουμε x_0 = seed

B2. Υπολογίζουμε το

$$x_i = \left(a x_{i-1}\right) \bmod m$$

και θέτουμε

$$U_i = \frac{x_i}{m}, i = 1, 2, \ldots, n$$

Γεννήτριες Τυχαίων Αριθμών

- **Πιθανότητα επανάληψης τιμών (όχι τυχαιότητα)**

- Επομένως, θα πρέπει να επιλέξουμε τα α, m έτσι ώστε για κάθε αρχική τιμή x_0, το πλήθος των βημάτων μέχρι την επανάληψη της διαδικασίας να είναι αρκετά μεγάλο. Γενικά τα α, m θα πρέπει να επιλεγούν έτσι ώστε:

Γεννήτριες Τυχαίων Αριθμών

1. Για κάθε αρχική τιμή x_0, η ακολουθία x_1/m, x_2/m, ... να μπορεί να θεωρηθεί ως μία πραγματοποίηση ακολουθίας ανεξάρτητων ομοιόμορφων στο(0,1) τυχαίων μεταβλητών.
2. Για κάθε αρχική τιμή x_0, το πλήθος των βημάτων μέχρι την επανάληψη της διαδικασίας να είναι αρκετά μεγάλο.
3. Οι αριθμοί x_n να παράγονται σχετικά εύκολα και γρήγορα από έναν υπολογιστή.

Γεννήτριες Τυχαίων Αριθμών

- x(1)=1; *\/δίνω την αρχική τιμή στο x*
- a=input ('dwse a\n'); *\/ζητάω να μου δοθεί από το πληκτρολόγιο τιμή για το a*
- m=input ('dwse m\n'); *\/ζητάω να μου δοθεί από το πληκτρολόγιο τιμή για το m*
- if (mod(m,4)==0) && (mod(a-1,4)==0) *\/ελέγχω αν τα a και m που μου έχουν δοθεί πληρούν τις προϋποθέσεις μου, δηλαδή αν το m διαιρείται με το 4 και αν το a-1 διαιρείται με το 4.*
- for i=2:20 *\/ξεκινάω μία επαναληπτική διαδικασία για δημιουργία X(i).*
-     x(i)=mod(a*x(i-1),m); *\/τύπος δημιουργίας X(i).*
- End *\/τέλος επαναληπτικής διαδικασίας.*
- for i=1:2 *\/ξεκινάω δεύτερη επαναληπτική διαδικασία για να βρω τα U(i) μου πού είναι οι τυχαίοι αριθμοί που ζητάω να βρω.*
-     u(i)=x(i)/m *\/διαιρώ τα X(i) με το m για να μου προκύψουν ψευδοτυχαίοι αριθμοί ανάμεσα στο 0 και στο 1.*
- End *\/τέλος δεύτερης επαναληπτικής διαδικασίας*
- *\/αν δεν πληρούνται οι προϋποθέσεις που θέλω για τα a και m τότε ζητάω από το χρήστη να ξανατρέξει το πρόγραμμα και να δώσει σωστά a και m.*
- else
-     disp('dwse allo m,a tetoia wste to m na diaireitai me to 4 kai to a-1 na diaireite me to 4\n')
- break;
-
- End *\/τέλος προγράμματος.*

Γεννήτριες Τυχαίων Αριθμών

- **μικτός τύπος**

$$x_{n+1} = (\alpha \cdot x_n + c) \bmod m$$

1. Για να έχουμε καλύτερα αποτελέσματα θα πρέπει το m να είναι δύναμη του 2, δηλαδή $m = 2^b$
2. Η μέγιστη περίοδος ψευδοτυχαίων που παράγονται με τον πολλαπλασιαστικό τύπο, όταν $m = 2^b$ και ισχύουν οι συνθήκες:

ι) $\alpha \bmod 4 = 1$ και

ιι) c και m πρώτοι αριθμοί μεταξύ τους,

είναι ίση με 2^b

Γεννήτριες Τυχαίων Αριθμών

Παράδειγμα:

Έστω $m = 8 = 2^3$, $\alpha = 5$ ($5 \bmod 4 = 1$), $c = 3$ και $x_0 = 3$, τότε έχουμε:

i	x_i	$5x_i + 3$	$5x_i+3 \bmod 8$
1	3	18	2
2	2	13	5
3	5	28	4
4	4	23	7
5	7	38	6
6	6	33	1
7	1	8	0
8	0	3	3
9	3		

18/8 = 2
υπόλοιπο 2

η περίοδος της παραγόμενης ακολουθίας των ψευδοτυχαίων είναι $8 = 2^3 = m$.

Γεννήτριες Τυχαίων Αριθμών

Εναλλακτικά :

$$U_i = \frac{x_i}{m}$$

$$x_i = \left(ax_{i-1} + c\right) \bmod m$$

x mod n συμβολίζει το υπόλοιπο της διαιρέσεως του x διά m

x_o είναι ακέραιος <m που λέγεται σπόρος (seed)

Οι αριθμοί x_i είναι ακέραιοι <m. Οι υπόλοιποι αριθμοί είναι ακέραιοι και ικανοποιούν τις σχέσεις

m > 0, α <m, και c <m.

Γεννήτριες Τυχαίων Αριθμών

η ακολουθία $\{x_i\}$ είναι περιοδική με περίοδο $p \leq m$. Οταν $p=m$ η γεννήτρια έχει **πλήρη περίοδο** και στην περίπτωση αυτή έχει μεγάλη ποικιλία διαφορετικών αριθμών x_i

Θεώρημα:

Η γεννήτρια $\{x_i\}$ έχει πλήρη περίοδο αν και μόνο εάν ισχύουν οι συνθήκες:

α. Ο μόνος φυσικός αριθμός που διαιρεί τα m και c είναι ο 1.

β. Αν q είναι κάποιος πρώτος αριθμός που διαιρεί το m, τότε διαιρεί και το α–1.

γ. Αν το 4 διαιρεί το m, τότε διαιρεί και το α–1.

Γεννήτριες Τυχαίων Αριθμών

Οταν c=0 έχουμε την περίπτωση των πολλαπλασιαστικών γεννητριών οι οποίες **δεν είναι πλήρους περιόδου.**

Θεώρημα:

Η γεννήτρια $\{x_i\}$ με c=0 έχει περίοδο p=m–1 αν ισχύουν:

α. Το m είναι πρώτος αριθμός.

β. Ο μικρότερος ακέραιος k για τον οποίον το $\alpha^k - 1$ διαιρείται από το m, είναι ο k=m–1.

Έλεγχοι τυχαιότητας

- **Ελεγχος ισοκατανομής.**

Χωρίζουμε το διάστημα (0,1) σε k ισομήκη υποδιαστήματα.

Παράγουμε n τυχαίους αριθμούς $U_1, \ldots, U_n$ και για κάθε υποδιάστημα j βρίσκουμε το πλήθος n_j των αριθμών που βρίσκονται σε αυτό.

$$X^2 = \frac{k}{n} \sum_{j=1}^{k} \left(n_j - \frac{n}{k} \right)^2$$

ακολουθεί προσεγγιστικά την κατανομή χ^2 με k–1 βαθμούς ελευθερίας.

Έλεγχοι τυχαιότητας

- Η υπόθεση ότι οι μεταβλητές U_1, ..., U_n είναι ομοιόμορφα κατανεμημένες στο διάστημα (0,1) είναι αποδεκτή με επίπεδο εμπιστοσύνης 1–α αν

$$X^2 \leq \chi^2_{\alpha,k-1}$$

Από πίνακες της κατανομης βρισκουμε το σημείο

$$P(X^2 \leq \chi^2_{\alpha,k-1}) = 1-\alpha.$$

Έλεγχοι τυχαιότητας

- **Έλεγχος συσχετίσεως.**

Παράγουμε n τυχαίους αριθμούς $U_1,\ldots, U_n$. Εξετάζουμε την ύπαρξη μη αποδεκτής συσχέτισης μεταξύ τους. Θεωρητικά η συσχέτιση τάξεως j

αμετάβλητης στοχαστικής διαδικασίας με ανεξάρτητα U_i είναι 0 για j>0 :

$$\rho_j = \frac{\mathrm{Cov}\{U_i, U_{i+j}\}}{\mathrm{Var}\{U_i\}} \ (\text{εξορισμού}) = \frac{E\{(U_i - E\{U_i\})(U_{i+j} - E\{U_{i+j}\})\}}{E\{U_i^{\,2}\} - [E\{U_i\}]^2}$$

$$= \frac{E\{U_i U_{i+j}\} - E\{U_i\}E\{U_{i+j}\}}{E\{U_i^{\,2}\} - [E\{U_i\}]^2} = \frac{E\{U_i\}E\{U_{i+j}\} - E\{U_i\}E\{U_{i+j}\}}{E\{U_i^{\,2}\} - [E\{U_i\}]^2}$$

$$= 0$$

Έλεγχοι τυχαιότητας

ενώ ρ_o=1. Ελέγχουμε λοιπόν αν οι συσχετίσεις της ακολουθίας που παράγεται $U_1, U_2, \ldots, U_n$ είναι ≠0.

Για ομοιόμορφα κατανεμημένες τ.μ. U_i η δειγματική συσχέτιση ρ_j τάξεως j υπολογίζεται ως εξής.

Διατρέχουμε την ακολουθία των αριθμών $U_1, U_2, \ldots, U_n$ με βήματα μήκους j αρχίζοντας από τον U_1 και έχουμε τους $U_1, U_{1+j}, U_{1+2j}, U_{1+3j} \ldots$.

$$\widehat{\rho}_j = -3 + \frac{12}{h+1} \sum_{k=0}^{h} U_{1+kj} U_{1+(k+1)j}$$

$$h = -1 + \mathrm{INT}[(n-1)/j].$$

Έλεγχοι τυχαιότητας

$$\mathrm{Var}\{\widehat{\rho}_j\} = \frac{13h+7}{(h+1)^2}.$$

Από το θεώρημα κεντρικού ορίου ισχύει

$$Z = \frac{\widehat{\rho}_j - 0}{\sqrt{\mathrm{Var}\{\widehat{\rho}_j\}}}$$

ακολουθεί προσεγγιστικά την κανονική κατανομή με μέση τιμή 0 και διασπορά 1.

Η υπόθεση ότι οι U_i είναι ασυσχέτιστες (τυχαίες) απορρίπτεται με πιθανότητα 1–α αν η Z έχει μεγάλο μέγεθος:

$$|Z| > z_{\alpha/2}. \quad \Rightarrow \quad P(Z > z_{\alpha/2}) = \alpha/2 \Leftrightarrow P(-z_{\alpha/2} < Z < z_{\alpha/2}) = 1-\alpha$$