Kas arvuti oskab üldse juhuslikku arvu välja mõelda?

1 kuu tagasi - 26.08.2026 Autor: AM

(Sisuturundus)

Kutsud välja funktsiooni *random()* ja saad numbri. Küsid uuesti, saad teise. Masina sees ei ole aga midagi juhuslikku: protsessor teeb sama sisendiga alati sama asja. Kust see number siis tuleb?

Algoritm, mis teeskleb

Pseudojuhuslik generaator on valem. Ta võtab ühe algväärtuse, mida nimetatakse seemneks, ja toodab sellest pika arvujada. Jada näeb välja juhuslik, olles samal ajal täielikult ette määratud: sama seemnega saad sama jada uuesti. See on kasulik omadus, sest nii saab simulatsiooni või vigase testi täpselt kordama panna.

Seemne valik on omaette lõks. Kui programm võtab selle kellaajast, saavad kaks samal sekundil käivitunud protsessi ühe ja sama jada. Sellist viga on aastate jooksul leitud kõikjalt, loosimängudest parooligeneraatoriteni. Sama kehtib liiga väikese seemne kohta: kui võimalikke algväärtusi on miljon, saab ründaja need miljon lihtsalt läbi proovida.

Pythoni dokumentatsioon on selle koha pealt ebatavaliselt otsekohene. Keel kasutab Mersenne Twisterit, mille periood on 2^19937−1, ja dokumentatsioon lisab kohe, et olles täiesti determineeritud, ei sobi see kõikideks eesmärkideks ning on täiesti sobimatu krüptograafiliseks kasutuseks.

Põhjus on see, et jada saab tagasi arvutada. Kui oled näinud 624 järjestikust 32-bitist väljundit, saad neist kokku panna generaatori sisemise oleku ja seejärel kõik järgnevad numbrid. Valmis tööriistad selleks on olemas ja keegi ei pea neid ise kirjutama.

Kust tuleb päris juhuslikkus

Päris juhuslikkus tuleb füüsikast. Operatsioonisüsteem kogub entroopiat kohtadest, kus maailm on korrapäratu: katkestuste ajastusest, ketta ja võrgu värinast, klaviatuuri mikroviivitustest. Uuemates protsessorites on selle jaoks eraldi käsud RDRAND ja RDSEED, mis võtavad müra otse kiibilt.

Kogutu läheb krüptograafilisse generaatorisse, mille väljundist sisemist olekut tagasi ei arvuta. Linuxis on need *getrandom()* ja */dev/urandom*, Pythonis moodul *secrets*. Reegel on lühike: kui kellegi teise õige oletus maksaks sulle raha või ligipääsu, ei ole tavaline *random()* see koht, kust numbrit võtta.

Nõrgad kohad ei istu tavaliselt serverites. Nad istuvad odavates seadmetes, mis on tehtud hinnale ja mitte turvalisusele, mis on ühtlasi üks põhjus läbi mõelda, kui palju eraldi koduvõrke sul tegelikult vaja on.

Praktikas käib rünnak muidugi mööda kõige lühemat teed. Petukõne on odavam kui ühegi generaatori ennustamine ja töötab paremini.

Kuidas juhuslikkust mõõta

Juhuslikkust ei saa tõestada. Saab otsida mustrit ja seda mitte leida, mis on nõrgem väide, kui esmapilgul tundub.

Standardsed testipatareid teevad täpselt seda. Tuntuim on NIST SP 800-22, kõrval käivad Dieharder ja TestU01. Nad vaatavad bittide sagedust, seeriaid, korrelatsioone, spektrit ja veel kümmet omadust.

Konks on selles, et Mersenne Twister läbib need testid hästi. Ta on üks põhjalikumalt testitud generaatoreid üldse ja samal ajal täiesti ennustatav. Test ütleb, et jada ei paista korrapärane. Ta ei ütle, et keegi ei suuda seda jätkata. Vastupidises suunas on testist rohkem kasu: kui generaator kukub juba sagedustestis läbi, ei pea edasi vaatama.

Sellepärast käib reguleeritud valdkondades testi kõrval alati kellegi allkiri. Õnnemängud on üks selline valdkond: Eestis annab korraldusloa Maksu- ja Tolliamet, kes peab ka avalikku nimekirja neist, kellel luba on.

Kus vahe maksab raha

Tavainimene puutub selle vahega kokku kahes kohas. Esimene on krüptovõtmed, mida ta kunagi ei näe. Teine on mängud, kus iga tulemus on generaatori tõmme.

Siin tuleb mängu number, mida enamik mängijaid on näinud ja vähesed lõpuni lugenud: tagastusprotsent ehk RTP. Kui mängu deklareeritud RTP on näiteks 96, tähendab see, et sajast panustatud eurost liigub keskmiselt 96 mängijatele tagasi. Ülejäänu on korraldaja eelis, mis on generaatori jaotusse sisse ehitatud.

Sõna „keskmiselt“ teeb kogu töö. RTP on kokkuvõte miljonitest tõmmetest ega ütle sinu õhtu kohta mitte midagi, täpselt nagu testi läbimine ei ütle midagi ennustatavuse kohta. Keskmine kirjeldab jaotust, mitte üksikjuhtu.

Testimislaborite töö on siin igav ja mahukas. Generaator pannakse tootma miljoneid tulemusi, jaotust võrreldakse deklareerituga ja jadast otsitakse ennustatavust. Sisuliselt sama, mida teeb NIST-i patarei, ainult et lõpus on nimi ja kuupäev.

Mängija jaoks on probleem selles, et sertifitseeritud ja sertifitseerimata generaator näevad ekraanil ühtemoodi välja. Kontrollida saab ainult paberit. Riiklik register küll ütleb, kellel luba on, aga sealt edasi läheb andmete otsimine tüütuks käsitööks. Et tarbija ei peaks ise ridades näpuga järge ajama, on Kasiinokompass selle eeltöö ära teinud: ametlikud registrikanded, boonustingimused ja väljamakseajad on pandud ühte mugavasse tabelisse. Loata korraldajate puhul polegi seal midagi kokku kanda, sest puudub algne garantii.

Inimene tahab ilusamat juhust

Kõige veidram osa loost on see, et päris juhuslikkus ei meeldi meile.

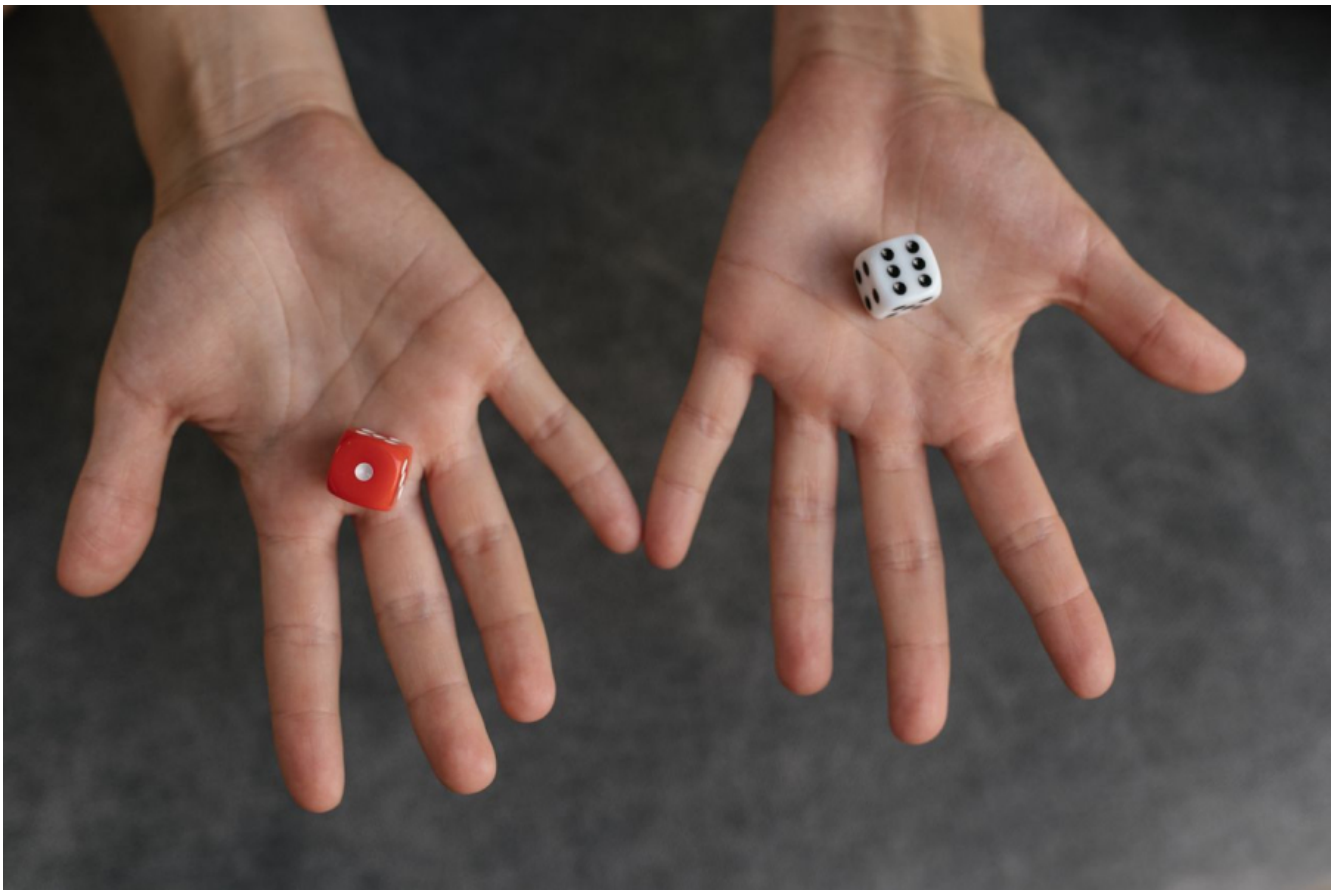
Spotify kirjeldas eelmisel aastal, kuidas matemaatiliselt korrektne segamine tekitas kaebusi. Kuulajad nägid järjestuses kordusi ja mustreid, sest juhuslikus jadas tulevad kobarad ette. Uus režiim tekitab mitu juhuslikku järjestust, hindab neid selle järgi, mida kasutaja on hiljuti kuulanud, ja valib värskeima. Insenerid sõnastasid selle nii, et statistiline juhuslikkus ei kandu alati üle tajutavaks juhuslikkuseks.

Sama ootus töötab mängija vastu. Kui viis korda järjest ei tulnud midagi, tundub kuuendal korral, et nüüd on aeg. Generaator ei tea, mitu korda sa oled kaotanud.

Kolm asja ühe sõna all

Arvutis tähendab „juhuslik" kolme eri asja: korratavat, ennustamatut ja meeldivat. Simulatsioon tahab esimest, turvalisus teist, kasutajaliides kolmandat. Enamik vigu selles vallas tuleb sellest, et võeti see juhuslikkus, mis oli käepärast, ja loodeti, et see on ka see, mida vaja oli.

- Uudised

- Sisuturundus

Pilt