General Test Procedures for Studying Random Data

A. "Chi-square" Test

E.g. "Throwing Two Dices"

s: value of the sum of two dices
p_s: probability

s =	2	3	4	5	6	7	8	9	10	11	12
p_s =	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{12}$	$\frac{1}{9}$	$\frac{5}{36}$	$\frac{1}{6}$	$\frac{5}{36}$	$\frac{1}{9}$	$\frac{1}{12}$	$\frac{1}{18}$	$\frac{1}{36}$

If we throw dices 144 times

	s =	2	3	4	5	6	7	8	9	10	11	12
observed #	Y_s =	2	4	10	12	22	29	21	15	14	9	6
expected #	np_s =	4	8	12	16	20	24	20	16	12	8	4

pair of dice is loaded?
We can't make a definite yes-or-no statement but we can give a probabilistic answer.

We form the statistic (Chi-Square Statistic)

$$V = \sum_{1 \le s \le k} \frac{(Y_s - np_s)^2}{np_s}$$

k: # of categories
n: # of observations

$$= \frac{1}{n} \sum_{1 \le s \le k} \left(\frac{Y_s^2}{p_s} \right) - n$$

$\nu = k - 1$: degree of the freedom

Table of CHI-SQUARE Distribution

Entry in row v under column p is x, which means

"The quantity V will be less than or equal to x with approximate probability p, if n is large enough."

Example:

value of s =	2	3	4	5	6	7	8	9	10	11	12
Experiment 1, Y_s =	4	10	10	13	20	18	18	11	13	14	13
Experiment 2, Y_s =	3	7	11	15	19	24	21	17	13	9	5

$V_1 = 29\frac{59}{120}$ $\qquad V_2 = 1\frac{11}{120}$

Discussion: V_1 is too high, V 0.1% of time

V_2 is too low, V 0.01% of time

Both represents a significant departure from randomness

To use chi-square distribution table: n should be large

How large should n be?

Rule of Thumb: n should be large enough to make each np_s be 5 or greater.

Chi-square Test:

1. Large number n of independent observations
2. Count the # of observations in k categories
3. Compute V
4. Look up Chi-Square distribution table

$< 1\%$ or $> 99\%$, reject

$1\% <$ $< 5\%$ or $95\% <$ $< 99\%$, suspect

$5\% <$ $< 10\%$ or $90\% <$ $< 95\%$, almost suspect

otherwise accept.

B. The Kolmogorov - Smirnov Test

Chi-square test: for discrete random data
KS test: for continuous random data

Def: $F(x)$ = probability that $(X \le x)$

n independent observations of the random quanity X

$X_1, X_2, \cdots, X_n$

Def: empirical distribution function $F_n(x)$

$$F_n(X) = \frac{\text{numbers of } X_1, X_2, \cdots, X_n \text{ that } \le x}{n}$$

KS Test is based on the difference btw $F(x)$ and $F_n(x)$

$$K_n^+ = \sqrt{n} \max_{-\infty < x < +\infty} (F_n(x) - F(x))$$

maximum deviation when F_n is greater than F

$$K_n^- = \sqrt{n} \max_{-\infty < x < +\infty} (F(x) - F_n(x))$$

maximum deviation when F_n is less than F

We get a similar table like the chi-square to find the percentile

Unlike χ^2, the table fits for any size of n

Simple procedure to obtain K_n^+, K_n^-

1. Obtain observations $X_1, X_2, \cdots, X_n$
2. Rearrange into ascending order
   $X_1 \le X_2 \le \cdots \le X_n$
3. $K_n^+ = \sqrt{n} \max_{1 \le j \le n} \left(\frac{j}{n} - F(X_j)\right)$

   $K_n^- = \sqrt{n} \max_{1 \le j \le n} \left(F(X_j) - \frac{j-1}{n}\right)$

Dilemma: we need large n to differentiate F_n and F
large n will average out local nonrandom behavior

Compromise: Take a moderate size of n, say 1000
Make a fair large number of K^+_{1000} on different part of random sequence
$K^+_{1000}(1), K^+_{1000}(2), \cdots, K^+_{1000}(r)$

Apply another KS test
the distribution of K_n^+ is approximated
$F_\infty(x) = 1 - e^{-2x^2}$

Significance: Detect both local & global nonrandom behavior

Empirical Tests: 10 tests

Test of real number sequence
$\langle U_n \rangle = U_0, U_1, U_2 \cdots$

Test of integer number sequence
$\langle Y_n \rangle = Y_0, Y_1, Y_2 \cdots$

$$Y_n = \lfloor d U_n \rfloor$$

Y_n: integers $[0, d-1]$

A. Equidistribution test (Frequency test)

Two ways:

① Use χ^2 test

d intervals

count the # of sequence $\langle Y_n \rangle = Y_0, Y_1, Y_2, \cdots$
falling into each interval

$k = d$

$p_s = 1/d$

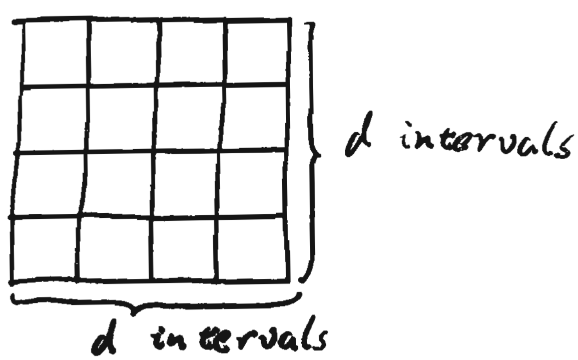
② Use KS test

test $\langle U_n \rangle = U_0, U_1, \cdots,$

$F(x) = x$ for $0 \le x < 1$

B. Serial Test

Pairs of successive numbers to be uniformly distributed



$k = d^2$

$p_s = 1/d^2$

Total: d^2 categories

Serial Test can be regarded as 2-D frequency test

Can be generalized to triples, quadruples, ...

C Gap test

Examine the length of "gaps" btw occurrence of U_j in a certain range

$$0 \le \alpha < \beta \le 1$$

gap: lengths of consecutive subsequences

$$U_j, U_{j+1}, \dots, U_{j+r}$$

lies btw α and β

Algorithm: 1. Initialize: $j \leftarrow -1$, $s \leftarrow 0$

2. $r \leftarrow 0$

3. $[\alpha \le U_j \le \beta]$?

$\begin{cases} \text{Yes} & j \leftarrow j+1 \\ \text{No} & \text{goto to } 5 \end{cases}$

4. $r \leftarrow r+1$, goto 3

5. record gap length

if $r \geq t$, $COUNT[t] \leftarrow COUNT[t]+1$

else $COUNT[r] \leftarrow COUNT[r]+1$

6. repeat until n gaps are found

$COUNT[0], COUNT[1], \dots, COUNT[t]$ should have the following probability

$$p_0 = p,\ p_1 = p(1-p),\ p_2 = p(1-p)^2, \dots, p_{t-1} = p(1-p)^{t-1}$$

$$p_t = (1-p)^t$$

$$p = \beta - \alpha$$

Now we can apply χ^2 test

Special cases: $(\alpha, \beta) = (0, \frac{1}{2}) \leftarrow$ runs above the mean

$(\alpha, \beta) = (\frac{1}{2}, 1) \leftarrow$ runs below the mean

D. Poker test (Partition test)

Consider 5 successive integers $(Y_{5j}, Y_{5j+1}, Y_{5j+2}, Y_{5j+3}, Y_{5j+4})$

Seven patterns

All different: abcde

One pair: aabcd

Two pairs: aabbc

Three of a kind: aaabc

Full house: aaabb

Four of a kind: aaaab

Five of a kind: aaaaa

Simplify:

5 different = all different

4 different = one pair

3 different = two pairs, or three of a kind

2 different = full house, or four of a kind

1 different = five of a kind

Generalized:

n groups of k successive numbers (k-tuples) with r different values

$$p_r = \frac{d(d-1)\cdots(d-r+1)}{d^k} \left\{ {k \atop r} \right\}$$

d: # of categories

Then χ^2-test can be applied

E. Coupon collector's test

In sequence $Y_0, Y_1, \dots$

the lengths of segments $Y_{j+1}, Y_{j+2}, \dots, Y_{j+r}$

to get a complete set of integers from 0 to $d-1$

Algorithm: 1. initialize: $j \leftarrow -1$, $s \leftarrow 0$, $COUNT[r] \leftarrow 0$ for $d \le r \le t$

2. $q \leftarrow r \leftarrow 0$, $OCCURS[k] \leftarrow 0$ for $0 \le k < d$

3. $r \leftarrow r+1$, $j \leftarrow j+1$

4. Complete Set? $OCCURS[Y_j] \leftarrow 1$ and $q \leftarrow q+1$

if $q = d$, a complete set

$q < d$, goto 3

5. [record the length]

$r \geq t$? $COUNT[t] \leftarrow COUNT[t]+1$: $COUNT[r] \leftarrow COUNT[r]+1$

6. UNTIL n values are found

Chi-square test can be applied to

$COUNT[d], COUNT[d+1], \dots, COUNT[t]$

$$p_r = \frac{d!}{d^r} \left\{ {r-1 \atop d-1} \right\}, \quad d \le r < t$$

$$p_t = 1 - \frac{d!}{d^{t-1}} \left\{ {t-1 \atop d} \right\}$$

F. Permutation test

A t-tuple $(U_{jt}, U_{jt+1}, \ldots, U_{jt+t-1})$ can have $t!$ possible relative orderings

For Example: $t = 3$

there should be $3! = 6$ categories

$1 < 2 < 3$	$2 < 1 < 3$	$2 < 3 < 1$
$1 < 3 < 2$	$3 < 1 < 2$	$3 < 2 < 1$

$k = t!$ $\quad p_s = 1/t!$

we can apply χ^2 test now

G. Run test

examine the length of monotone subsequences

"runs up" increasing

"runs down" decreasing

$$| \underbrace{1\ 2\ 9}_{3} | \underbrace{8}_{1} | \underbrace{5}_{1} | \underbrace{3\ 6\ 7}_{3} | \underbrace{0\ 4}_{2} |$$

COUNT[i]

Note: χ^2 test cannot be directly applied because of lack of independence (each segment depends on previous segment)

Then we need to calculate

$$V = \frac{1}{n} \sum_{1 \le i, j \le 6} (COUNT[i] - nb_i)(COUNT[j] - nb_j)\, a_{ij}$$

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} & a_{56} \\ a_{61} & a_{62} & a_{63} & a_{64} & a_{65} & a_{66} \end{pmatrix} = \begin{pmatrix} 4529.4 & 9044.9 & 13568 & 18091 & 22615 & 27892 \\ 9044.9 & 18097 & 27139 & 36187 & 45234 & 55789 \\ 13568 & 27139 & 40721 & 54281 & 67852 & 83685 \\ 10891 & 36187 & 54281 & 72414 & 90470 & 111580 \\ 22615 & 45234 & 67852 & 90470 & 113262 & 139476 \\ 27892 & 55789 & 83685 & 111580 & 139476 & 172860 \end{pmatrix}$$

$$(b_1\ b_2\ b_3\ b_4\ b_5\ b_6) = \left(\tfrac{1}{6}\ \tfrac{5}{24}\ \tfrac{11}{120}\ \tfrac{19}{720}\ \tfrac{29}{5040}\ \tfrac{1}{840}\right)$$

Then V should have the χ^2 distribution with degree 6

H. Maximum-of-t test

Examine the maximum value

let $V_j = \max(U_{tj}, U_{tj+1}, \ldots, U_{tj+t-1})$

The distribution $F(x) = x^t$

Then we can apply the KS test here

I. Collision test

Suppose we have m urns and n balls

$m >> n$

most of the ball will fall in an empty urn

if a ball falls in an urn that already has a ball

we call it "collision"

A generator passes the collision test only if it doesn't induce too many or too few collisions.

Let $m :=$ # of urns

$n :=$ # of balls

Probability c collisions occur:

$$\frac{m(m-1)\cdots(m-n+c+1)}{m^n}\left\{ {n \atop n-c} \right\}$$

J. Serial correlation test

Consider the observations $(U_0, U_1, \cdots, U_{n-1})$

and $(U_1, \cdots, U_{n-1}, U_0)$

test the corelation btw these two tuples

We compute:

$$C = \frac{n(U_0U_1 + U_1U_2 + \cdots + U_{n-2}U_{n-1} + U_{n-1}U_0) - (U_0 + U_1 + \cdots + U_{n-1})^2}{n(U_0^2 + U_1^2 + \cdots + U_{n-1}^2) - (U_0 + U_1 + \cdots + U_{n-1})^2}$$

A "good" C should be btw $\mu_n - 2\sigma_n$ and $\mu_n + 2\sigma_n$

$$\mu_n = \frac{-1}{n-1}, \quad \sigma_n = \frac{1}{n-1}\sqrt{\frac{n(n-3)}{n+1}}, \quad n > 2$$

The Spectral Test

Ideas underlying the test.

Congruential generators generate random numbers in grid!

In t-dimensional space

$\{(U_n, U_{n+1}, \cdots, U_{n+t-1})\}$

Idea: Compute the distance btw lines (2D), planes (3D), parallel hyperplane (>3D)

$1/\nu_2$: maximum distance btw lines
two dimensional accuracy

$1/\nu_3$: maximum distance btw planes
three dimensional accuracy

$1/\nu_t$: maximum distance btw hyperplanes
t-dimensional accuracy

Difference btw truly random sequences and periodic sequences

truly random sequences: accuracy remains same in all dimension

periodic sequences: accuracy decreases when $t \uparrow$

Spectral test is by far the most powerful test

All "good" generators pass it.

All known "bad" generators fail it.

Summary

1. Basic idea of empirical tests
   the combination of random numbers is expected
   to conform to a specific distribution

   $\Downarrow$

   ① Build the combination
   ② Use χ^2 or KS test to test the deviation
   from the expected distribution

2. We can perform ∞ # of tests

3. We might be able to construct a test to "kill"
   a specific generator

Other resources of test

1. FFT, metropolis, wolfgang tests (spectrum)
2. Diehard (http://www.stat.fsu.edu/~Diehard)
3. SPRNG (implements most of the emperical tests
   and spectrum tests)
   http://sprng.cs.fsu.edu