Lecture 12 Brownian motion, chi-square distribution, d.f.

- Adjusted schedule ahead
- Chi-square distribution (lot of supplementary material, come to class!!!) 1 lecture
- Hypothesis testing (about the SD of measurement error)and P-value (why n-1?supplement) 1 lecture
- Chi-square test for Model validation (chapter 11)
- Probability calculation (chapter 4)
- Binomial distribution and Poisson (chapter 5, supplement, horse-kick death cavalier data, hitting lottery, SARS infection)
- Correlation, prediction, regression (supplement)
- t-distribution, F-distribution

Brownian motion-random zigzag movement of small particles dispersed in fluid medium, R. Brown (1773-1853, Brit. Botanist)

Molecule movement is unpredictable

Position along each coordinate axis is modeled as a random variable with normal distribution (like pollen in water)

2- D case. A particle moves randomly on the glass surface, starting from (0, 0) position. The position one minute later is at (X,Y). Suppose EX=0, EY=0, SD(X)=SD(Y)=1 mm. Find the probability

that the particle is within 2mm of the original.

How about within 3mm? Or in general within c mm? That is, what is Pr(distance< c)?

How big a circle has be drawn in order to have 95% chance of containing the particle?

$X \sim$ normal(0,1), $Y \sim$normal(0,1)

independent

Pythagorean theorem: right triangle : $c^2 = a^2 + b^2$

Distance between two points on a plane with coordinates (a, b) and (c.d) is equal to square root of $(a-c)^2 + (b-d)^2$; give examples

Squared distance between (X,Y) to (0,0) is

$D^2=X^2+Y^2$ (why?)

The name of the distribution of this random variable is known as a chi-square distribution with 2 degrees of freedom. (denoted by ~ chisq, df=2; or χ^2 d.f=2; or $\sim\chi^2_2$) ; computer programs are available for

Finding Pr $(X^2 + Y^2 < a)$ for any positive value a. But Table on page 567 gives something different. It gives information on the upper tail, Pr($X^2 + Y^2 > a$).

- Use simulation :
- Suppose there are 1000 particles released at

The origin and move independent of each other; one minute later, record their locations . Mark locations; Draw histogram of squared distance from the origin

Find the proportion of particles that are within 1mm, 2 mm, 3 mm, and so on..

How to use table of chi-square distribution on page 567

- At the first column, find df =2
- The row corresponding to df is relevant to chi-quared distribution with two degrees of freedom
- Find the value 13.82 at the rightmost end
- Look up for the column header : .001
- This means that there is only a probability of .001 that the chi-squared random variable will be greater than 18.82

```
> (sqrt 13.82)
3.7175260590882213
> (sqrt 10.60)
3.255764119219941
> (sqrt 9.210)
3.034798181098704
> (sqrt 7.378)
2.716247411411156
> (sqrt 5.991)
2.4476519360399265
> (sqrt 4.605)
2.1459263733874936
> (sqrt 3.794)
1.947819293466414
> (sqrt 3.219)
1.7941571837495174
```

Movement in three dimension space

A point with Coordinate (X, Y, Z);
squared distance from origin is
$D^2 = X^2 + Y^2 + Z^2$ If X, Y,Z are independent , normal (

Then D^2 follows a Chi-square distribution with three degrees of freedom

In general, movement in n dimension space; a point with coordinate (X_1, X_2, ..X_n); squared distance from origin is $D^2 = X_1^2 + X_2^2 + \ldots + X_n^2$ If X_1, ..X_n are independent, normal (0,1), then D^2 follows a Chi-square distribution with n degrees of freedom istance of a point (x,y) from diagonal line :

ook at the symmetric point (y,x)

he projection must be at the middle; so squared distance is

$$R^2 = (X - C)^2 + (Y-C)^2, \text{ where } C=(X+Y)/2$$

Follows a chi-squared with one degree of freedom

Losing one degree of freedom due to projection constraint (fr
another viewpoint, one equation (X-C) + (Y-C)=0 to hold;)

-Dimension case

Projection to the diagonal (x=y=z) line

$$R^2 = (X-C)^2 + (Y-C)^2 + (Z-C)^2; C= (X+Y+Z)/3$$

Follows a chi-square distribution with (3-1)=2 degrees of freedo

-dimensional case :

$(X_1-C)^2 + (X_2-C)^2 + \ldots + (X_n-C)^2$; $C = (X_1 + .. + X_n)/n$ =average

ollows a chi-square distribution with n-1 degrees of freedom

If variance of normal each X is c

- Then D^2/σ^2 follows a chi-square distribution with n degrees of freedom
- R^2/σ^2 follows a chi-square distribution with n-1 degrees of freedom ; this is also true even if the mean of the normal distribution (for each X) is not zero (why?)