Computer Systems Modelling

[Relates to goodness of fit tests, simulation of queueing systems & validating simulations.]

(a) A goodness of fit can be used to help validate a modelling assumption that a sample of random variables comes from a given probability distribution, in this case from the distribution $P(X=i) = 1/6$ $i = 1, 2, \ldots, 6$.
Under the null hypothesis that the sample arises from the supposed distribution we construct a test statistic T, say and reject the null hypothesis if $T > t$ where

$$P(T > t) = 0.05, \text{ say.}$$

Different rejection regions can be used by replacing 0.05 by 0.01, for example.

(b) The χ^2 test is constructed as follows
Set N_i = # outcomes with value i
$i = 1, \ldots, 6$

$n = 6$, size of state space
Define test statistic,

$$T = \sum_{i=1}^{n} \frac{(N_i - 150 \cdot 1/6)^2}{150 \cdot 1/6}$$

where $150 \cdot \frac{1}{6}$ is the expected number of outcomes of any type in a sample of size 150.

(b) ctd.

Under the null hypothesis that $P(X=i) = 1/6$ $i=1,\ldots,6$

$$T \sim \chi^2_{n-1} \quad \text{distribution}$$

with $n-1=5$ degrees of freedom (at least, asymptotically for large sample sizes (& 150 is certainly large))

The rejection region is $T > t$ where t is given by

$P(T > t) = 0.05$, say, with T given by the null hypothesis.

Applying the χ^2 test to the given sample yields

$$T = \sum_{i=1}^{n} \frac{(N_i - 25)^2}{25}$$

$$= \frac{(3^2 + 4^2 + 3^2 + 2^2 + 3^2 + 11^2)}{25}$$

$$= \frac{168}{25}$$

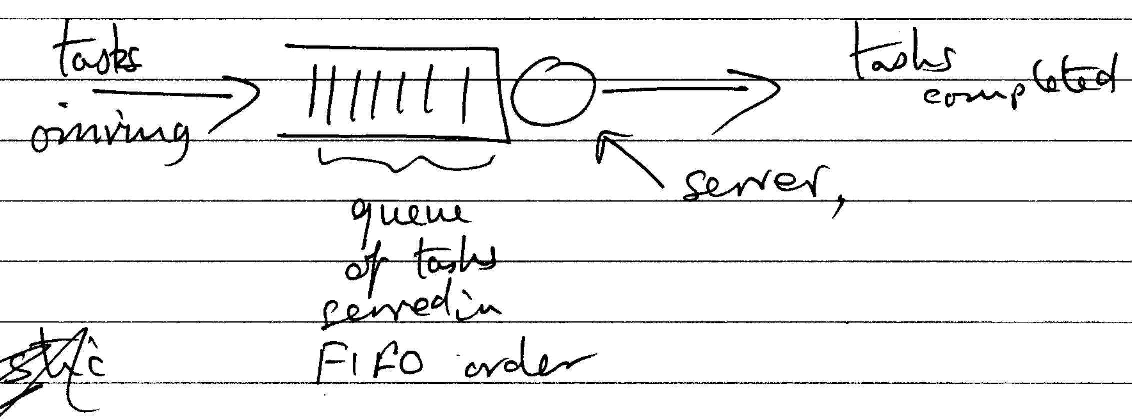
$$= 6.72$$

From the supplied table

$$P(T > 11.07) = 0.05$$

and since $T = 6.72 < 11.07$ we <u>donot</u> reject the null hypothesis and are content that the random number generator is unbiased.

(c) FIFO M/G/1 queue



~~Probabilistic~~ ~~Arrival~~

Probabilistic modelling assumptions:

(1) Interarrivals times are independent $Exp(\lambda)$ some fixed, λ, with mean rate of arrivals given by λ.
Corresponds ~~to~~ Poisson process

(2) tasks have service times given by some given distribution with distribution function $G(\cdot)$.

(3) tasks are served, in order of arrival with the next task to be served entering service immediately upon arrival if the server is empty or immediately upon the currently executing task completing.

Log events A_i $i=1,\ldots$
= time of i^{th} arrival
S_i = time of i^{th} task entering service
C_i = time of i^{th} task completing ~~leaving~~ service

Define

$$X_i = A_{i+1} - A_{i}$$ $i = 1, \ldots,$

as the i^{th} interarrival time

$$Y_i = C_i - S_i$$ $i = 1, \ldots,$

as the i^{th} service time.

Then observe ~~each~~ system and use Kolmorgorov - Smirnov goodness of fit test to assess assumptions

X_i is sample from $Exp(\lambda)$ where λ is the chosen arrival rate

Y_i is sample from $G(\cdot)$ the given service distribution.

The Kolmorgorov - Smirnov test is the the equivalent of the χ^2 test for continuous distributions instead of discrete distributions.

Letting

$$F_e(x) = \frac{\#\text{ obs in sample} \leq x}{n}$$

The test statistic is

$$D = \max_{x} \left| F_e(x) - F(x) \right|$$

where $F(\cdot)$ is the distribution for the null hypothesis (either $Exp(\lambda)$ or $G(\cdot)$ depending on arrivals ~~and~~ service.)

The distribution of D under the null hypothesis has been tabulated for varying sample size n.