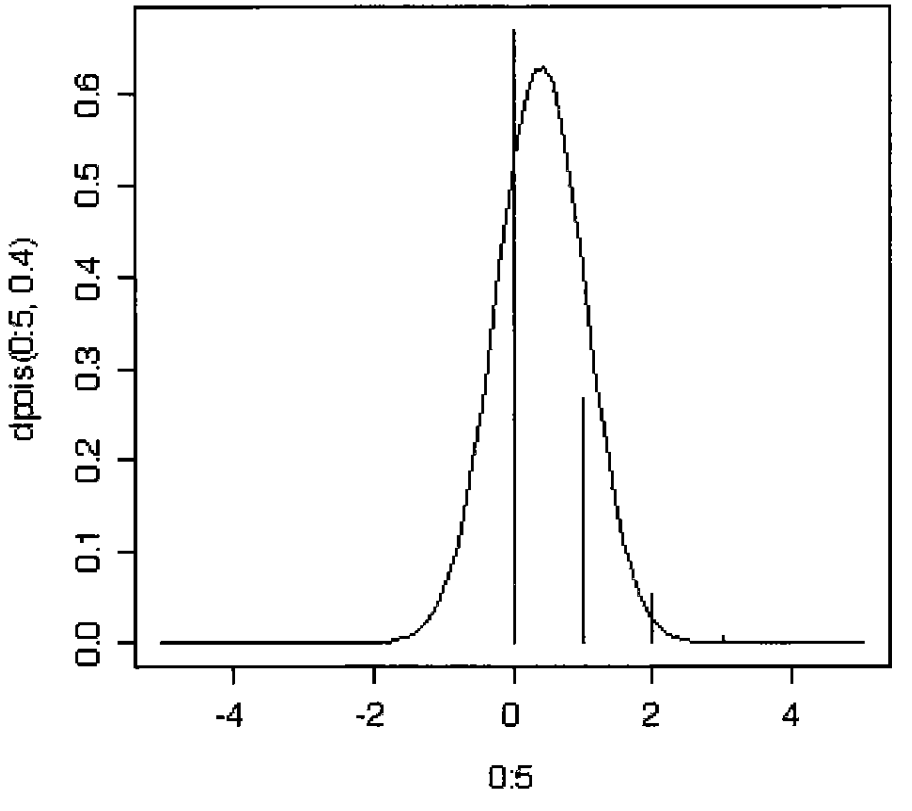
STATISTICS 3N03/3J04 2004-11-01

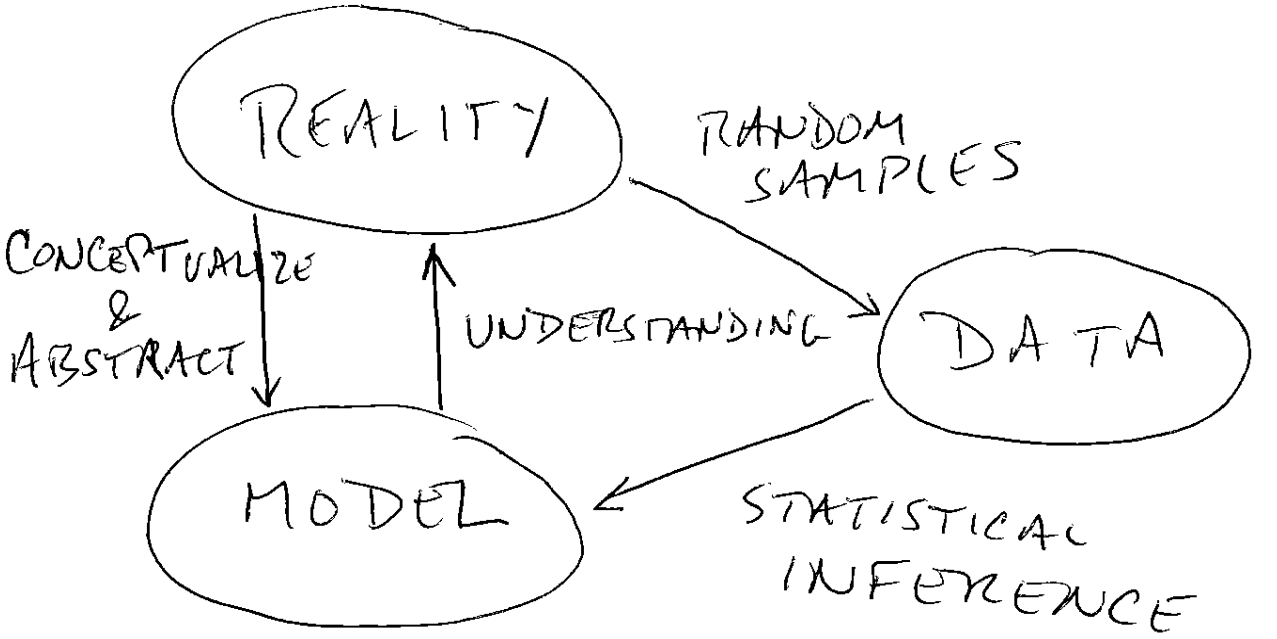
20-1

Notes for Exercise 4-71 p. 122

```
> pp <- 1-ppois(0, .4)
> pp
[1] 0.3296800
> 1 - pnorm((.5-.4)/sqrt(.4))
[1] 0.4371835
> 1 - pnorm((0-.4)/sqrt(.4))
[1] 0.7364554
> 1 - pbinom(350, 1000, pp)
[1] 0.08119912
> 1 - pnorm((350.5 - 1000*pp)/sqrt(1000*pp*(1-pp)))
[1] 0.08067642
> 1 - pnorm((350 - 1000*pp)/sqrt(1000*pp*(1-pp)))
[1] 0.08582804
> plot(0:5, dpois(0:5, .4), type="h", xlim=c(-5,5))
> xgr <- seq(-5,5,len=80)
> lines(xgr, dnorm(xgr, .4, sqrt(.4)), type="l", col="red")
```



STATISTICAL INFERENCE



FREQUENTIST INFERENCE

- PARAMETERS OF THE MODEL ARE FIXED BUT UNKNOWN

- STATISTICS FOLLOW A SAMPLING DISTRIBUTION UNDER REPEATED SAMPLING FROM THE POPULATION

- OUR UNCERTAINTY ABOUT THE PARAMETER COMES FROM THE SAMPLING DISTRIBUTION OF THE STATISTIC

- METHODS OF INFERENCE:
  - P-VALUE
  - TEST OF HYPOTHESIS AT A GIVEN LEVEL OF SIGNIFICANCE
  - CONFIDENCE INTERVAL

EX $X_1, \dots, X_n \sim IN(\mu, \sigma^2)$

μ: UNKNOWN OF INTEREST

σ^2: KNOWN, NOT OF INTEREST

$\Rightarrow \bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$

SAMPLING DISTRIBUTION SAYS THAT WITH PROBABILITY 95%

$$\mu - 1.96\frac{\sigma}{\sqrt{n}} < \bar{X} < \mu + 1.96\frac{\sigma}{\sqrt{n}}$$

$\bar{X}$: ↑ RANDOM

μ: FIXED BUT UNKNOWN

AFTER $\bar{x}$ IS OBSERVED, WE ARE 95% CONFIDENT THAT

$$\bar{x} - 1.96\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + 1.96\frac{\sigma}{\sqrt{n}}$$

μ: FIXED BUT UNKNOWN

$\bar{x}$: OBSERVED

i.e. THE RANDOM INTERVAL

$$\bar{X} - 1.96\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + 1.96\frac{\sigma}{\sqrt{n}}$$

WILL COVER THE TRUE VALUE OF μ 95% OF ALL SAMPLES SO WE ARE 95% CONFIDENT IT DOES FOR THIS SAMPLE.

EX $\underset{\sim}{x} = (23.1, 27.2, 24.4, 25.0)$

GIVEN: $\sigma = 1.2$

"WE ARE 95% CONFIDENT THAT THE TRUE MEAN LIES IN THE INTERVAL (23.75, 26.10)"

PROBABILITY INTERPRETATION:

IF WE REPEATED THIS MANY TIMES WITH NEW, INDEPENDENT SAMPLES, WE WOULD GET A DIFFERENT INTERVAL EVERY TIME, AND 95% OF THEM WOULD COVER THE TRUE MEAN μ.

NOTE:

- REFERS TO SAMPLES YOU DIDN'T GET!
- μ IS FIXED BY REALITY BUT UNKNOWN
- $\bar{x}$ IS FIXED BY WHAT YOU OBSERVED

SO: WE SAY "CONFIDENCE" INSTEAD OF "PROBABILITY".

TEST OF HYPOTHESIS

(9.2.1 p. 289)

Ex SAME DATA. THE STANDARD SAYS THAT THE MEAN SHOULD BE 30. WE OBSERVED $\bar{x} = 24.925$. IS THIS CONSISTENT WITH $\mu = 30$?

μ TRUE MEAN - UNKNOWN

μ_0 HYPOTHESIZED MEAN, FROM THEORY OR FROM A STANDARD

$\bar{X}$ SAMPLE MEAN AS A RANDOM VARIABLE BEFORE YOU OBSERVE IT.

$\bar{x}$ OBSERVED SAMPLE MEAN

σ "NUISANCE PARAMETER" GIVEN IN THIS EXAMPLE.