2004-10-20

15-1

STATS 3N03/3J04

MAKEUP TESTS

FRI NOV 5 & 26

9:30 - 10:30 REF - 102

ERRATA IN TEXT

- ON WILEY WEB SITE
- LINK FROM COURSE PAGE

DEFINITIONS TO LOOK UP AND THINK ABOUT

COEFFICIENT PARAMETER

TWO PARAMETERIZATIONS OF THE EXPONENTIAL DISTRIBUTION

$$f(x|\lambda) = \lambda e^{-\lambda x} \qquad x > 0$$

$$f(x|\mu) = \frac{1}{\mu} e^{-\frac{x}{\mu}} \qquad x > 0$$

λ RATE PARAMETER

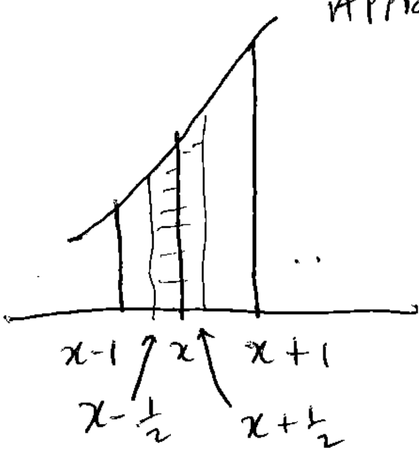
$\mu = \frac{1}{\lambda}$ MEAN OR SCALE PARAMETER

NORMAL APPROXIMATION TO THE BINOMIAL

$$X \sim BIN(n, p)$$

$\Rightarrow$ IF n IS LARGE AND p IS NOT TOO CLOSE TO 0 OR 1,

$$X \underset{\text{APPROX}}{\sim} N(np, np(1-p))$$



$$P(X = x) \approx \underset{N(np,\, npq)}{P} \left(x - \tfrac{1}{2} < X < x + \tfrac{1}{2}\right)$$

$$P(X \le x) \approx \Phi\left(\frac{x + \frac{1}{2} - np}{\sqrt{npq}}\right)$$

CONTINUITY CORRECTION

EX PROBABILITY OF GETTING ≤ 3 HEADS IN 10 TOSSES OF A FAIR COIN

NORMAL APPROXIMATION TO THE POISSON

$$X \sim POIS(\mu)$$

$\Rightarrow$ IF μ IS LARGE,

$$P(X \le x) \approx \Phi\left(\frac{x+\frac{1}{2}-\mu}{\sqrt{\mu}}\right)$$

CONTINUITY CORRECTION:
USE <u>ONLY</u> WHEN APPROXIMATING A DISCRETE DISTRIBUTION BY A CONTINUOUS DISTRIBUTION.

<u>EX</u> ON AVERAGE, 400 CARS/DAY PASS A COUNTER. WHAT IS THE PROBABILITY OF GETTING MORE THAN 450 CARS ON A GIVEN DAY. STATE ANY ASSUMPTIONS AND DISCUSS THEIR VALIDITY.

1 - ppois(450, 400)

0.00652

1 - pnorm(450.5, 400, 20)

0.00578

$$P(X > 450) = 1 - P(X \le 450)$$

$$\approx 1 - \Phi\left(\frac{450.5 - 400}{\sqrt{400}}\right)$$

$$= 1 - \Phi\left(\frac{50.5}{20}\right) = 1 - \Phi(2.525)$$

$$= 1 - 0.9942 = 0.0058$$

PROBABILITY PLOTS TO COMPARE TWO DISTRIBUTIONS

→ "QUANTILE - QUANTILE" OR "Q-Q" PLOTS TO TEST FOR NORMALITY

PLOT QUANTILES OF DATA AGAINST QUANTILES OF STANDARD NORMAL;

IF DATA CAME FROM A NORMAL DISTRIBUTION GET A STRAIGHT LINE

INTERCEPT = MEAN

SLOPE = STD DEVIATION

DATA: $x_1, \dots, x_n$

SORTED DATA: $x_{(1)}, \dots, x_{(n)}$

$$x_{(1)} \le x_{(2)} \le \dots \le x_{(n)}$$

CUMULATIVE DISTRIBUTION FUNCTION

$$F(x) = P(X \le x)$$

SAMPLE ESTIMATE AT $x_{(i)}$

$$\hat{F}(x_{(i)}) \approx \frac{i - .5}{n}$$

CASE: NORMAL DISTRIBUTION

$$F(x_{(i)}) = \Phi\left(\frac{x_{(i)} - \mu}{\sigma}\right)$$

SOLVE FOR $x_{(i)}$:

$$x_{(i)} = \mu + \sigma\,\Phi^{-1}(F(x_{(i)}))$$

↑ STANDARD NORMAL QUANTILE FUNCTION

OBSERVED: $x_{(i)}$

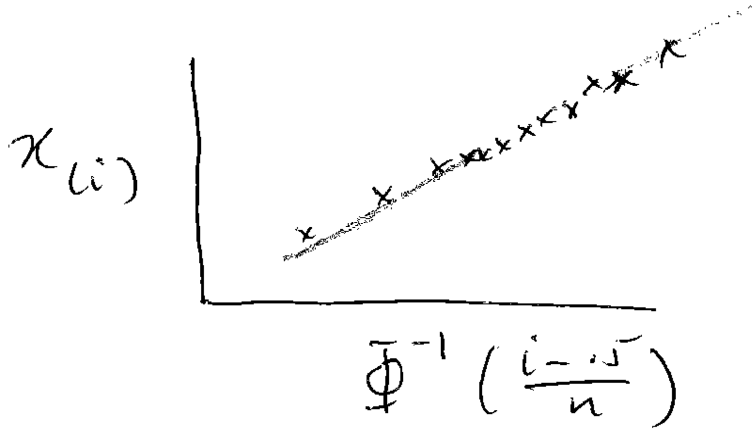
EXPECTED IF NORMAL:

$$\mu + \sigma\,\Phi^{-1}(F(x_{(i)}))$$

ESTIMATE OF EXPECTED IF NORMAL

$$\mu + \sigma\,\Phi^{-1}(\hat{F}(x_{(i)}))$$

$$= \mu + \sigma\,\Phi^{-1}\left(\frac{i - .5}{n}\right)$$




```
n <- length(x)
plot(qnorm(((1:n)-.5)/n), sort(x))
```

OR:

```
qqnorm(x)
qqline(x)
```

↖ ADDS A STRAIGHT LINE THROUGH FIRST AND THIRD QUARTILES.

HOW STRAIGHT SHOULD THE LINE BE?

- NON-DECREASING POINTS
- POINTS SPARSE IN TAILS
- LOOK AT SIMULATIONS