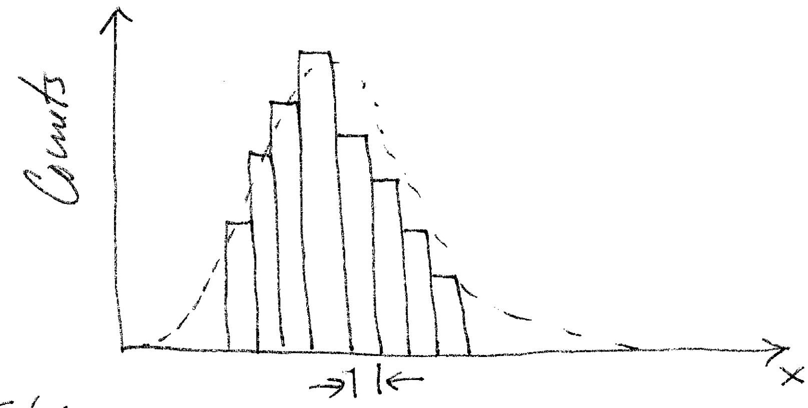
Chi-Square Test of a Distribution

Conditions:

- N measurements x_i; $i=1,\dots,N$ of a variable x
- Define bins of width Δ_j
- Construct histogram
- Need to determine how well $p(x)$ predicts the data

Note 1: In each bin $\Delta_j \in [x_j, x_{j+1}]$, the number of counts of x should be $N\,p(x_j)$

Note 2: The counts in each bin should be thought of as a Poisson variable.



J bins Δ centered on x_j

Observed number of counts in bin Δ_j is n_j

Thus we write the error function:

$$\chi^2 \equiv \sum_{j=1}^{J} \frac{[n_j - N\,p(x_j)]^2}{\sigma_j^2}$$

- By construction, we expect χ^2 to be a number of order J
- Numerator is observed spread in obs - theory
- Denominator is expected spread

In reality, $\langle \chi^2 \rangle = \nu = J - J_c$

where ν = # of degrees of freedom

J = # of sampled bins

J_c = # of constraints

A more common form of χ^2 is

$$\chi^2 = \sum_{j=1}^{J} \frac{\left[n_j - N\,p(x_j) \right]^2}{N\,p(x_j)}$$

$$\cong \sum_{j=1}^{J} \frac{\left[n_j - N\,p(x_j) \right]^2}{n_j}$$

Since for Poisson variables,

$$\sigma_j^2 = N\,p(x_j) \cong n_j$$

i.e., variance = mean

- In general, if $N p(x)$ is chosen independently of n_j, we have $\nu = J - 1$, since we have constrained the variance $\sigma_j^2 = N p(x)$

- Usually we work with reduced χ_ν^2

$$\chi_\nu^2 = \chi^2 / \nu$$

So we expect $\langle \chi_\nu^2 \rangle \to 1$ in general

- The variation of χ^2 is described by the χ^2 distribution.

- The survivor function of the χ^2 distribution is called the p-statistic

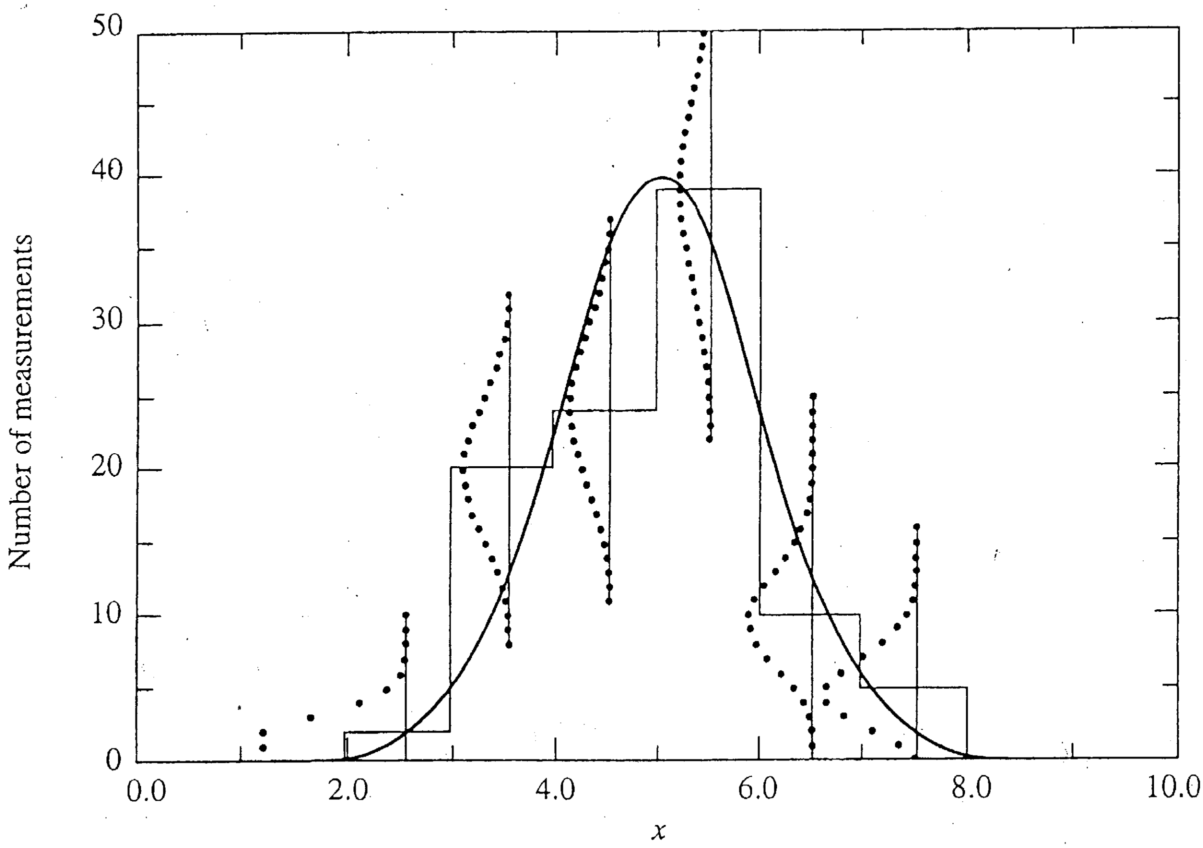


FIGURE 4.1
Histogram, drawn from a Gaussian distribution mean $\mu = 5.0$ and standard deviation $\sigma = 1$, corresponding to 100 total measurements. The parent distribution $y(x_j) = NP(x_j)$ is illustrated by the large Gaussian curve. The smaller dotted curves represent the Poisson distribution of events in each bin, based on the sample data.