1.7.7 Estimating Distributions (continued).

Lec 23 Wed 10/17/01

Last time (Lab 7):

Ex 2. Observe $X(1), \dots, X(N)$ – iid, pdf $f(x)$.

key:
- independent experiments
- all have same pdf

To estimate $p^{(i)} = \int_{x_{i-1}}^{x_i} f(x)\,dx$,

$$\hat{p}_N^{(i)} = \frac{\#\text{ outcomes in } [x_{i-1}, x_i]}{N}$$

$$E(\hat{p}_N^{(i)}) = p^{(i)}$$

$$\mathrm{Var}(\hat{p}_N^{(i)}) = \frac{p^{(i)} - (p^{(i)})^2}{N}$$

If $f(x) \approx$ const on $[x_{i-1}, x_i]$, then $\int_{x_{i-1}}^{x_i} f(x)\,dx \approx f(x)(x_i - x_{i-1})$

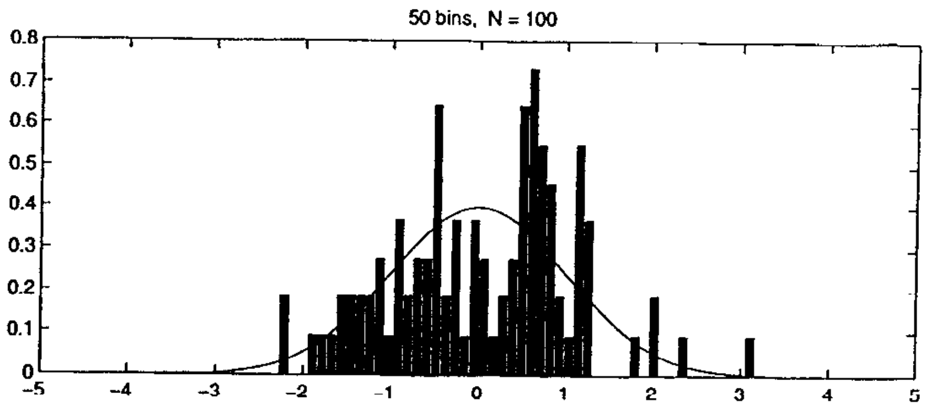
$\Rightarrow$ estimate $f(x)$ as

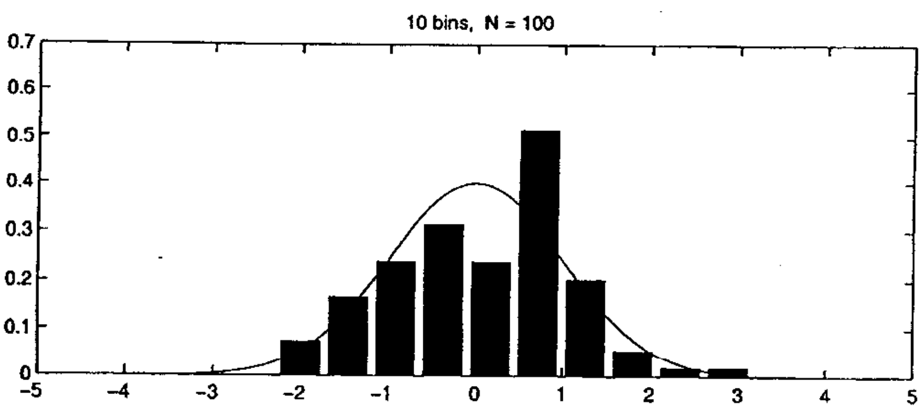
$$\hat{f}_N^{(i)} = \frac{\hat{p}_N^{(i)}}{x_i - x_{i-1}} \quad \text{on } [x_{i-1}, x_i].$$

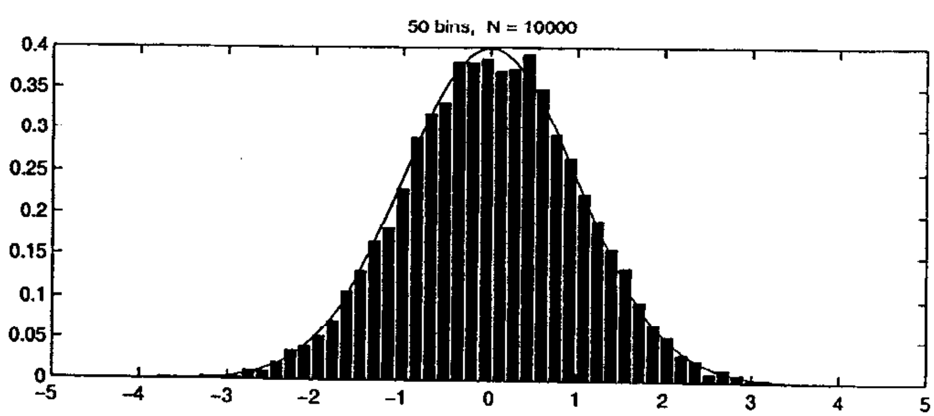
$$\mathrm{Var}(\hat{f}_N^{(i)}) = \frac{p^{(i)} - (p^{(i)})^2}{N(x_i - x_{i-1})^2}$$

- $x_i - x_{i-1}$ too small $\Rightarrow$ need large number of experiments N to get a reliable estimate
- $x_i - x_{i-1}$ too large $\Rightarrow$ $f(x)$ will not be const on $[x_{i-1}, x_i]$.

[this is illustrated on the next page]


50 bins, N = 100
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
−5 −4 −3 −2 −1 0 1 2 3 4 5


10 bins, N = 100
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
−5 −4 −3 −2 −1 0 1 2 3 4 5


50 bins, N = 10000
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
−5 −4 −3 −2 −1 0 1 2 3 4 5

Sometimes it is not necessary to estimate the whole pdf; often, there are practical situations, when there are only one or several parameters of interest.

Example 3. [Suppose we are interested in measuring some quantity A, but what we actually measure is:]

$$\underbrace{Y(1)}_{\text{what we actually observe}} = \underset{\substack{\uparrow \\ \text{the quantity of interest}}}{A} + \underbrace{X(1)}_{\text{measurement noise}}$$

$Y(2) = A + X(2)$ – 2nd measurement, independently obtained

$\vdots$

$Y(N) = A + X(N)$

Suppose it is known that the pdf of $X(n)$ is the same for all n, and $E(X(n)) = 0$, but the specific form of the pdf is unknown. How to estimate A?

One possibility: $\hat{A}_N = \frac{1}{N}\sum_{n=1}^{N} Y(n)$.

Then $E(\hat{A}_N) = \frac{1}{N}\sum_{n=1}^{N} E(Y(n)) = \frac{1}{N}\sum_{n=1}^{N} E(A + X(n))$

$= A$ – unbiased

$$\mathrm{Var}(\hat{A}_N) = \frac{1}{N^2}\cdot N\,\mathrm{Var}(X(n)) = \frac{\mathrm{Var}(X(n))}{N} \to 0 \text{ as } N \to \infty$$

(if $\mathrm{Var}(X(n)) < \infty$)

– consistent

The variance approaches zero as the number of measurements increases to ∞.

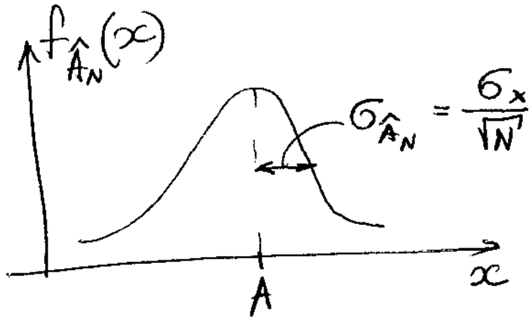
Central limit theorem:

Suppose $Z_N = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} X(n)$, where $X(n)$ are iid with mean zero and variance σ^2.

Then, as $N \to \infty$,
cdf of $Z_N \to$ Gaussian cdf with mean zero and var. σ^2

[The pdf of a Gaussian (normal) r.v. R with mean m and variance σ^2 is

$$f_R(r) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(r-m)^2}{2\sigma^2}}$$
]



- for large N, the distribution of our estimate ~~can be~~ is approx. Gaussian
- ~~for~~ as N gets larger, it becomes more and more probable that $\hat{A}_N$ ~~that~~ we compute from our observations is near A.

Example 4. How to estimate the variance of iid r.v.'s $X(1), \ldots, X(N)$?
(Also Lab 7)

Solution. (a) If the mean m is known,

try $$\hat{\lambda} = \frac{\sum_{n=1}^{N} \left(X(n) - m\right)^2}{N}$$

$$E(\hat{\lambda}) = \frac{1}{N} \sum_{n=1}^{N} E\left\{\left(X(n) - m\right)^2\right\} = \mathrm{Var}\left(X(n)\right) \Rightarrow \underline{\text{unbiased}}.$$

(b) If the mean m is unknown,

- estimate the mean as in Ex 3
$$\hat{m} = \frac{1}{N} \sum_{n=1}^{N} X(n)$$
- try
$$\hat{\lambda} = \frac{1}{N} \sum_{n=1}^{N} \left(X(n) - \hat{m}\right)^2$$
$$= \frac{1}{N} \sum_{n=1}^{N} \left(\left(X(n) - m\right) - \left(\hat{m} - m\right)\right)^2$$
$$= \frac{1}{N} \sum_{n=1}^{N} \left(X(n) - m\right)^2 - \underbrace{2\left(\hat{m} - m\right) \underbrace{\frac{1}{N} \sum_{n=1}^{N} \left(X(n) - m\right)}_{(\hat{m} - m)}}_{2(\hat{m} - m)^2} + \underbrace{\frac{1}{N} \sum_{n=1}^{N} \left(\hat{m} - m\right)^2}_{(\hat{m} - m)^2}$$
$$= \frac{1}{N} \sum_{n=1}^{N} \left(X(n) - m\right)^2 - \left(\hat{m} - m\right)^2$$

$$E(\hat{\lambda}) = \mathrm{Var}\left(X(n)\right) - \mathrm{Var}(\hat{m}) =$$
$$= \mathrm{Var}\left(X(n)\right) - \frac{\mathrm{Var}\left(X(n)\right)}{N} = \frac{N-1}{N} \mathrm{Var}\left(X(n)\right) - \text{biased!}$$

$$\hat{\lambda}' = \frac{1}{N-1} \sum_{n=1}^{N} \left(X(n) - \hat{m}\right)^2 \text{ is unbiased, because}$$

$$\hat{\lambda}' = \frac{N}{N-1} \hat{\lambda} \Rightarrow E(\hat{\lambda}') = \frac{N}{N-1} E(\hat{\lambda}) = \mathrm{Var}\left(X(n)\right).$$