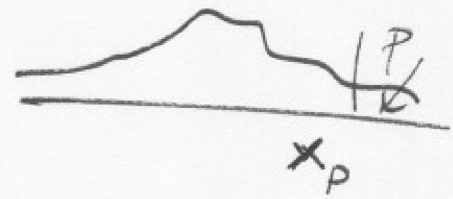
Ma 26
Winter 2008

CONFIDENCE INTERVALS BASED ON SAMPLE MEDIAN

Let $X_1, \dots, X_n \overset{iid}{\sim} f$

Let x_p = upper p^{th} quantile



Then
$$p = P(X > x_p) = \int_{x_p}^{\infty} f(x)\,dx$$

and
$$\underbrace{\sum_{i=1}^{n} 1\{X_i > x_p\}}_{\text{(Number of X's } > x_p)} \sim \text{Binomial}(n, p)$$

Assume n is odd. Let M_n = sample median.

Then
$$P(M_n > x_p) = P\left((\text{Number of } X\text{'s} > x_p) \text{ is } > \tfrac{n}{2}\right)$$
$$= P\left(\text{Binomial}(n,p) \text{ rv is } > \tfrac{n}{2}\right).$$

<u>Idea</u>: Given α, solve (1) $\boxed{P(\text{Bin}(n,p) > \tfrac{n}{2}) = \tfrac{\alpha}{2} \text{ for } p}$.

Then $P(M_n > x_p) = \frac{\alpha}{2} \Rightarrow \boxed{M_n \le x_p \text{ with prob. } 1 - \frac{\alpha}{2}}$

So the sample median gives a $100(1-\frac{\alpha}{2})\%$ confidence lower bound on the p^{th} quantile, x_p, for that p.

<u>Note</u>: This holds for any f ("non-parametrically").

<u>Application</u>: If $X_1, \dots, X_n \overset{iid}{\sim} f(x-\theta)$ where θ is unkown but f is <u>Known</u>, then $p = \int_{x_p}^{\infty} f(x-\theta)\,dx = \int_{x_p-\theta}^{\infty} f(x)\,dx$

$\Rightarrow x_p - \theta = C(p)$, where (2) $\boxed{\int_{C(p)}^{\infty} f(x)\,dx = p}$.

Rx: Choose α, solve (1) for p, then (2) for $C(p)$.

Then with probability $1-\frac{\alpha}{2}$, $M_n \le x_p = \theta + C(p)$, so $M_n - C(p)$ is a $100(1-\frac{\alpha}{2})\%$ confidence lower bound on θ.

If f is symmetrical about 0, $M_n \pm C(p)$ gives a $100(1-\alpha)\%$ CI for θ.