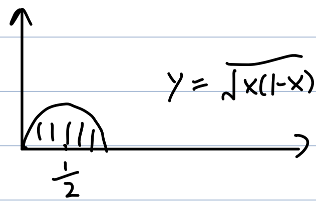
$$\frac{1}{2}\int_0^1 x\sqrt{1-x^2}\,x\,dx = \frac{1}{2}\int_0^1 \sqrt{u}\sqrt{1-u}\,du = \frac{1}{2}B\left(\frac{3}{2},\frac{3}{2}\right) = \frac{\pi}{16}$$

$u = x^2$, $x = \sqrt{u}$



$y = \sqrt{x(1-x)}$ $y^2 = x - x^2$ $y > 0$

Statistics vs Statistic

- Collecting -> design of experiment -> sample -> independent random sample (X1, … , Xn i.i.d on X)
  From population (X - random variable)
- Processing (descriptive / inferential)
- Analyzing (estimation)
- Interpreting (hypothesis test vs decision making)
- Presenting

Math

Data - plural

Glivenko-Cantelli Theorem (1933)

X1, … , Xn — i.i.d on X

F_X — cdf of X

$$\underbrace{\bar{F}_n(x)}_{\text{empirical cdf}} = \frac{1}{n}\sum_{k=1}^{n} I(X_k \le x)$$

$$I(A) = \begin{cases} 1 & \text{true} \\ 0 & \text{false} \end{cases}$$

$$P\left(\lim_{n\to\infty} \sup_{-\infty<x<\infty} |F_X(x) - \bar{F}_n(x)| = 0\right) = 1$$

fix x $\bar{F}_n(x) = \frac{1}{n}\sum_{k=1}^{n} I(X_k \le x) \xrightarrow{a.s.} F_X(x)$

$I(X_k \le x)$: i.i.d $B(1, F_X(x))$

$$P(I(X_k \le x) = 1) = P(X_k \le x) = F_X(x)$$

$$\sqrt{n}\,(F_n(x) - \bar{F}_n(x)) \xrightarrow{d} N(0, F_X(x)(1 - F_X(x)))$$

Order Statistics

Sample $X_1, \cdots, X_n \longrightarrow X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n-1)} \leq X_{(n)}$

X - continuous $\rightarrow X_{(1)} < X_{(2)} < \cdots < X_{(n)}$

with prob 1, no ties between two sample

$P(X_{(1)} > t) = P(X_1 > t, X_2 > t, \cdots, X_n > t)$

indep $\rightarrow = P(X_1 > t)\, P(X_2 > t) \cdots P(X_n > t)$

samp cdf $\rightarrow = (1 - F_X(t))^n$

$P(X_1 \leq t) = 1 - P(X_1 > t) = 1 - (1 - F_X(t))^n$

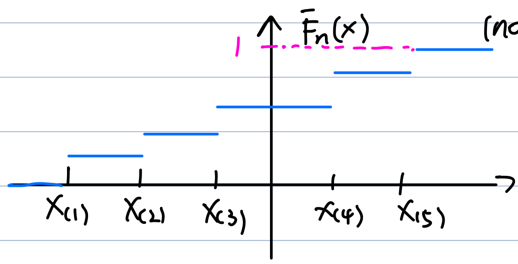
if x has pdf $f(x)$

$F'_X(t) = f(t)$

Th $X_{(1)}$ has pdf $f_{(1)}(t) = n\,(1 - F_X(t))^{n-1} f(t)$

similarly $P(X_{(n)} \leq t) = (F_X(t))^n$

pdf $\Rightarrow n\, f(t)\, (F_X(t))^{n-1}$ is pdf of $X_{(n)}$



(no ties) $\Leftarrow x_1 \neq x_2 \cdots \neq x_s$ $P(X_k = X_j) = 0$ $k \neq j$

empirical cdf