Probability Theory & Statistics

Innopolis University, BS-I,II

Spring Semester 2016-17

Lecturer: Nikolay Shilov

Part I

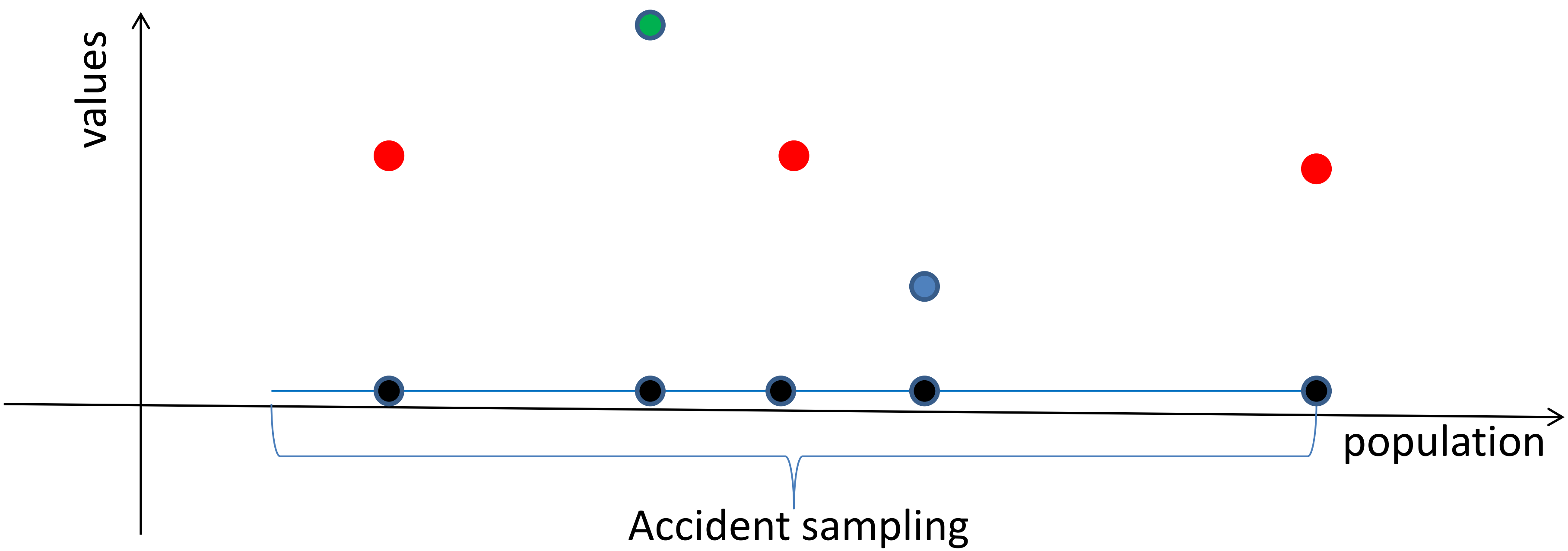
INTRO TO MATHEMATICAL STATISTICS

What is Mathematical Statistics?

- Mathematical Statistics is a branch of Probability Theory that studies infinite series of independent identically distributed (IID) random variables.

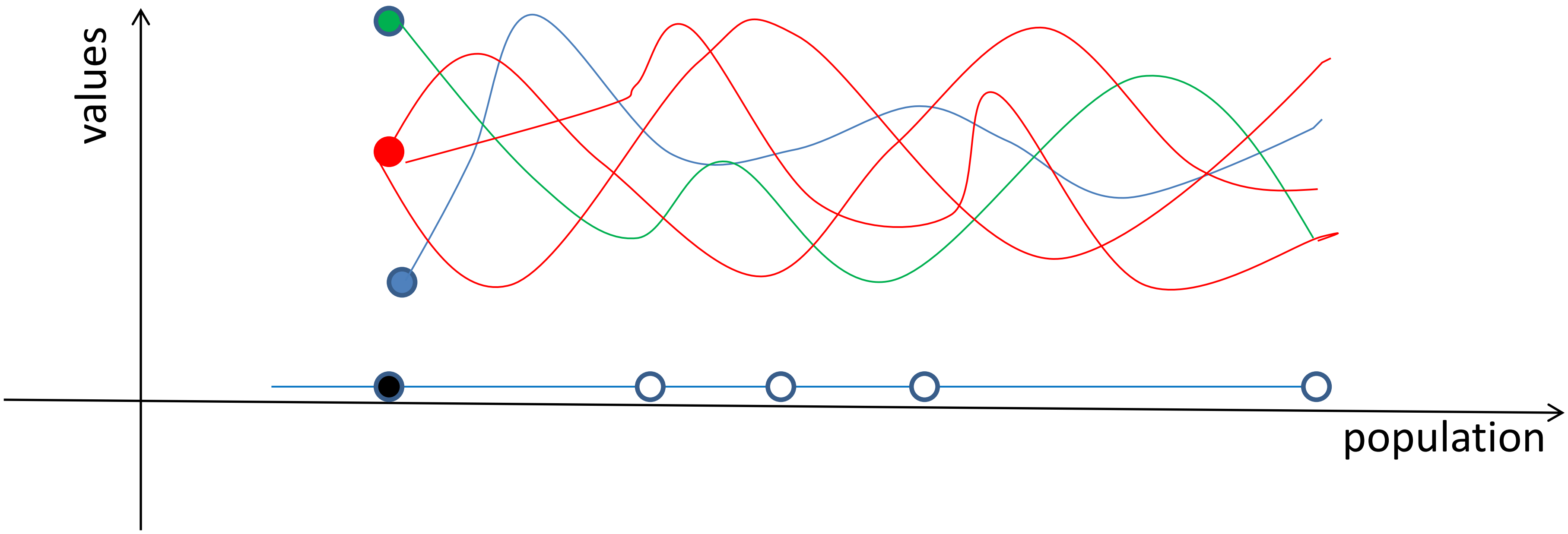
Non-probability sampling

Accidental (or *convenience*) *sampling* of size n is a sample drawn from that part of the *population* that is close to hand.



Mathematical random sampling

A *random sample* of length/size n for a (fixed) distribution F is a set of n IID random variables with distribution F.



Sample statistic

- A (*sample*) *statistic* is a single numeric measure of some attribute of a sample (e.g., its arithmetic mean value).
- More formally: a statistic is a function of a sample where the function is independent of the sample's distribution.
- The term statistic is used both for the function and for the value of the function on a given sample.

Statistics as Estimators

- An estimator is a rule to calculate an estimate a parameter (e.g. mean, variance) based on observed data.
- Estimator examples: estimate
  - the mean as the arithmetic mean in a random point $[X_1(r) + \ldots + X_n(r)]/n$;
  - the second moment as the mean squared in a random point $[(X_1(r))^2 + \ldots + (X_n(r))^2]/n$.

Consistent estimator

- An estimator of parameter is *consistent* if it *converges in probability* (due to use of a random population point) the true value.
- According to the week Law of Large Numbers (LLN, see lecture for 11)
  - the arithmetic mean is a consistent estimator for the mean,
  - the mean squared is a consistent estimator the for the second moment.

Exercise

- Explain (prove) that the mean squared is a consistent estimator the for the second moment.

Bias of a consistent estimator

- The *bias* of an estimator $T(X_1,\ldots X_n)$ of a parameter θ is the difference $E[T(X_1,\ldots X_n)] - \theta$.
- The bias of
  - the arithmetic mean estimation for the mean,
  - the mean squared estimation for the second moment

is 0.

Biased and Unbiased Estimators

- An estimator is *unbiased* if it has 0 bias. (Refer the previous slide for examples.)
- Any unbiased estimator + 1/n is a biased one, e.g.: the arithmetic mean + 1/n is a consistent biased estimator for the mean.
- An unbiased estimator may be non-consistent, e.g. (exercise): let estimator $T(X_1, \ldots X_n)$ be equal to the value of X1 in a random point.

Example a biased and unbiased estimations for the variance

- An intuitive estimator

$$\underline{\mathbf{D}}_n = \Sigma_{1\leq k\leq n}(X_k - \underline{\mathbf{X}}_n)^2/n = \Sigma_{1\leq k\leq n}X_k{}^2/n - \underline{\mathbf{X}}_n{}^2$$

  is consistent but bias:

$$E(\underline{\mathbf{D}}_n) = \sigma^2 - \sigma^2/n = [(n-1)/n]\,\sigma^2;$$

- in contrast estimator

$$\underline{\mathbf{D}}'_n = \Sigma_{1\leq k\leq n}(X_k - \underline{\mathbf{X}}_n)^2/(n-1)$$

  is unbiased and consistent.

Part II

CONFIDENCE

Confidence interval and level

- A *confidence interval* of a *confidence level* p ($0 \leq p \leq 1$) for a parameter θ is an interval [L,U] such that $P(L \leq \theta \leq U) = p$.
- Example - 68–95–99.7 rule (lecture for week 13): for μ
  - $[\mu-\sigma, \mu+\sigma]$ has level 68.27%;
  - $[\mu-2\sigma, \mu+2\sigma]$ has level 95.45%;
  - $[\mu-3\sigma, \mu+3\sigma]$ has level 99.73%.

Exercises

- Build confidence intervals for μ with levels σ, 2σ and 3σ for
  - uniform distribution on [0,1];
  - Bernoulli distribution with $p=2/3$.

Using Chebyshev's inequality

- Recall from lecture for week 11: if X is a random variable with a finite expectation μ and finite non-zero deviation σ, then

$$P(|X - \mu| \geq k\sigma) \leq 1/k^2$$

for any real number $k>0$.

Using Chebyshev's inequality (cont.)

- It implies (see proof of the weak LLN in lecture for week 11)
  - that $P(|\underline{\mathbf{X}}_n - \mu| \geq \varepsilon) \leq \sigma^2/(n\varepsilon^2)$
  - and $P(|\underline{\mathbf{X}}_n - \mu| < \varepsilon) \geq 1 - D(\underline{\mathbf{X}}_n)/\varepsilon^2$

  any $\varepsilon>0$.

Using Chebyshev's inequality (cont.)

- For a given confidence level p
  - compute the arithmetic mean m_n and variance d_n in a random point for $X_1, \ldots X_n$;
  - compute
    - $\varepsilon = [d_n/(1-p)]^{1/2}$ and
    - the confidence interval $[m_n - \varepsilon, m_n + \varepsilon]$ for the mean μ.

Bernoulli trials

- Recall that if $X_1,\ldots X_n$=Bernoulli(p) then
  - $E(\underline{\mathbf{X}}_n)=p$ and
  - $D(\underline{\mathbf{X}}_n)=p*q/n$;
- according to Moivre-Laplace / Central Limit theorem (for big n)

$$P\left(|\, p-\underline{X}_n \,|<\varepsilon\sqrt{pq/n}\right)\approx 2\Phi_0(\varepsilon)$$

Bernoulli trials (cont.)

- For a given confidence level t
  - compute the arithmetic mean p_n, its compliment q_n =1- p_n, and "diviation" d_n= $[p_n * q_n / n]^{1/2}$ in a random point for $\underline{\mathbf{X}}_n$;
  - compute ε such that $2\Phi_0(\varepsilon)$=t and the confidence interval $[p_n - \varepsilon * d_n, p_n + \varepsilon * d_n]$ for the mean p.

Part III

PROBABILISTIC INEQUALITIES

Less than by probability

- Let P be the probability of joint distribution of random variables X and Y; let $X<Y$ by probability, if $P(X<Y) > P(X \geq Y)$ (i.e. $P(X<Y) > ½$).
- Exercises:
  - give example(s) of $X<Y$ by probability;
  - is “less than by probability” a transitive relation?

Stochastically less than

- Let X and Y be random variables; let X<Y stochastically, if
- $P_X(X<z) \geq P_Y(Y<z)$ for all $z \in R$,
- and $P_X(X<z) > P_Y(Y<z)$ for some $z \in R$.
- Exercises:
  - give example(s) of X<Y stochastically;
  - is "stochastically less than" a transitive relation?

Exercise

- Let X be a random variable uniformly distributed on [0,1], and let Y be

$$\begin{cases} \varepsilon^2 * X \text{ with probability } \varepsilon; \\ X + \varepsilon^2 * (1 - X) \text{ with probability } 1-\varepsilon \end{cases}$$

for some fixed small $1>\varepsilon>0$.

- Show that Y>X with some probability p, but Y<X stochastically. (What exactly is this probability p?)

Confidence and inequalities

- Question: What type of probabilistic inequalities can/may be used to evaluate confidence level?