120AB Study Guide

mrevanishere

January 7, 2021

1 Notation

Note: finite or countably infinite is just countable (see infinities in proofsbook)
Notation: uppercase letter such as Y to denote an rv and a lowercase letter such as y to denote a particular value of that rv.
$(Y = y)$ is the set of all points in S assigned to the value y by rv Y.
$P(Y = y)$ is the probability that Y takes on the value y, defined as the sum of th proba of all sp in S that are assigned to value y. sometimes denoted by $p(y)$
...

2 Probability

DeMorgan's Laws: $\overline{A \cap B} = \overline{A} \cup \overline{B}$ and $\overline{A \cup B} = \overline{A} \cap \overline{B}$
Distributive laws:

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

Experiment: process by which observation is made
Event: E (simplest possible outcome)
SS (Sample Space): set of all possible sample points (sp) S of an experiment
dss (Discrete Sample Space): contains either a countable number of distinct sp
Mutually exclusive (ME) sets are ME events.
Compound Events: unions of sets of sp of simple events
A simple event E_i is included in event A iff A occurs whenever E_i occurs.
Event (dss): collection of sample points (any subset of S)

Relative Frequency Definition 2.6:
S is an SS with an experiment. To every event A in S a number P(A) the probability of A such that:

Axiom 1: $P(A) \geq 0$
Axiom 2: $P(S) = 1$
Axiom 3: If $A_{1-...}$ form a sequence of pairwise ME events in S $(A_i \cap A_j = \emptyset, if i \neq j) then$

$$P(A_1 \cap A_2 \cap A_3 \cap ...) j \sum_{i=1}^{\infty} P(A_i)$$

Sample-Point Method (2.5)
1.
2.
3.
4.
5.
Multiplication Principle (Fundamental Rule of Counting) (mn rule)
Permutations:

$$P_r^n = P(n, r) = seeproofsbook = \frac{n!}{(n-r)!}$$

Partitions: n objects into k groups containing n1-k objects where each object appears exactly in one group $\sum_{i=1}^{k} n_i = n$ is

$$N = \binom{n}{n1-k} = \frac{n!}{n1-k!}$$

see multinomials (2.6)
Combinations: n choose r

$$C_r^n = \binom{n}{r} = \frac{P_r^n}{r!} = \frac{n!}{r!(n-r)!}$$

Bayes' Theorem ()Conditional Probability (2.7)): probability of A given B

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Independent Events: iff

$$P(A|B) = P(A)$$
$$P(B|A) = P(B)$$
$$P(A \cap B) = P(A)P(B)$$

Multiplicative Law of Probability: intersection of events (and)

$$\begin{aligned} P(A \cup B) &= P(A)P(B|A) \\ &= P(B)P(A|B) \\ \text{if independent } P(A \cap B) &= P(A)P(B) \end{aligned}$$

Additive Law of Probability: The probability of the union of two events A and B is

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
$$\text{ME events } P(A \cup B) = P(A) + P(B)$$

$P(A) = 1 - P(\overline{A})$
Event Composition:
1.
2.
3.
4.
(HELP) Law of Total Probability (2.10): For some pos int k, let the sets B_{1-k} be such that
1. $S = B_1 \cup .... \cup B_k$
2. $B_i \cap B_j = \emptyset$, for $i \neq j$.
Then the collection B_{1-k} is a partition of S.
such that $P(B_i) > 0$, for $i = 1 - k$ Then for any event A:

$$P(A) = \sum_{i=1}^{k} P(A|B_i)P(B_i)$$

rv (Random Variable): real-valued function for which the domain is a SS.
Population: larger body of data where samples are taken from
(HELP) Replacement:
srs (simple random sampling): N and n is population and sample. A ssrs is each of the $\binom{N}{n}$ samples has an equal probability of being selected, the sampling is said to be random.

3 Discrete

Discrete - rv Y is discrete if can assume a countable number of distinct values
pd (Probability Distribution) - collection of probabilities
pf, pmf (Probability Function, pmf) for Y - $p(y)$
pd for drv Y that shows $p(y)$ for all y. Must be:
1. $0 \leq p(y) \leq 1$ for all y
2. $\sum_y p(y) = 1$, where summation is over all y with nonzero $p(y)$.
params (Parameters) - numerical descriptive measures for $p(y)$.
ev (Expected Value) for drv:

$$E(Y) = \sum_y yp(y)$$

if $p(y)$ is accurate of population frequency distribution then $E(Y) = \mu$ the population mean.
ev of g(Y) a real-valued funciton of Y:

$$E[g(Y)] = \sum_{\text{all } y} g(y)p(y)$$

variance of a drv Y with mean $E(Y) = \mu$: the ev of $(Y - \mu)^2$:

$$V(Y) = E[(Y - \mu)^2]$$

sd of drv Y: positive square root of V(Y)
if p(y) is accurate for population then $V(Y) = \sigma^2$, sd is σ.
Theorems (closed under addition and scalar multiplication):

$$E(c) = c$$
$$E[cg(Y)] = cE[g(Y)]$$
$$E[g_1(Y) + ... + g_k(Y)] = E[g_1(Y)] + ... + E(g_k(Y)])$$
$$V(Y) = \sigma^2 = E[(Y - \mu)^2] = E(Y^2) - \mu^2$$

3.1 binomial pd

Binomial Experiment:
1. consists of a fixed n identical trials

2. binary outcome S success, F failure
3. proba of S on a single trial is p and remains same from trial to trail. The proba of F is $q = 1 - p$
4. The trials are independent
5. rv Y, the number of S during n trials.
pf of drv Y is a binomial pd based on n trials with sucess proba p iff

$$p(y) = \binom{n}{y} p^y q^{n-y}, \ , y = 0 - n \text{ and } 0 \leq p \leq 1$$

see binomial theorem in proofs text
mean and variance of binomial drv (107):

$$\mu = E(Y) = np \text{ and } \sigma^2 = V(Y) = npq$$

note: $\sum p(y) = 1$, $E(Y^2 - Y) = E(Y(Y-1))$

3.2 geometric pd

Geometric Experiment: same as Binomial except that it is the number of the trial which the first success occurs
pdf of drv Y is said to have a geometric pd iff:

$$p(y) = q^{y-1}p, \, y = 1 - ..., 0 \leq p \leq 1$$

mean and variance of geometric drv:

$$\mu = E(Y) = \frac{1}{p} \text{ and } \sigma^2 = V(Y) = \frac{1-p}{p^2}$$

3.3 Negative Binomial pd

...

3.4 (HELP)Hypergeometric pd

??? $b = N - r$, SS method, multiplication principle
pdf of drv Y has a hypergrometric pd iff:

$$p(y) = \frac{\binom{r}{y}\binom{N-r}{n-y}}{\binom{N}{n}}$$

where y is an int 0-n, subject to $y \leq r$ and $n - y \leq N - r$.
fact: $\sum_{i=0}^{n} \binom{r}{i}\binom{N-r}{n-i} = \binom{N}{n}$
mean and variance of hypergeometric drv (127):

$$\mu = E(Y) = \frac{nr}{N} \text{ and } \sigma^2 = V(Y) = n\binom{r}{N}\binom{N-r}{N}\binom{N-n}{N-1}$$

see 128 semi proof

3.5 Poisson pd

$\lambda = np$ and take the limit of binomial pf to infinity (131)
pf of drv Y has a Poisson pd iff:

$$p(y) = \frac{\lambda^y}{y!}e^{-\lambda},\ y = 1 - \lambda > 0$$

mean and variance of Poisson drv (134)

$$\mu = E(Y) = \lambda \text{ and } \sigma^2 = V(Y) = \lambda$$

Poisson process: λ is mean num of occurrences per unit, then Y = the number of occurences in a units has a Poisson pd with mean $a\lambda$.

3.6 Moments and mgf

kth moment origin: kth moment of rv Y taken about origin is: $E(Y^k)$ denoted as μ'_k
kth moment mean: kth central moment of Y is $E[Y - u]^k$ denoted as μ_k where $\sigma^2 = \mu_2$
mgf (moment-generating function): for rv Y is $m(t) = E(e^{tY})$. mgf exists if there is a positive b such that m(t) is finite for $|t| \leq b$
T3.12 if m(t) exists, then for any pos int k

$$\frac{d^k m(t)}{dt^k}|_{t=0} = m^{(k)}(0) = \mu'_k$$

...

3.7 pgf

3.8 Tchebysheff

4 Continuous

df, cdf (cumulative distribution function): $F(y) = P(Y \leq y)$ for $-\infty < y < \infty$
cdf for drv are always step functions.
cdf properties: 1. $F(-\infty) \equiv \lim_{y\to-\infty} F(y) = 0$
2. $F(\infty) \equiv \lim_{y\to\infty} F(y) = 1$
3. $F(y)$ is a nondecreasing function of y meaning for any $y_1 < y_2$, then $F(y_1) \leq F(y_2)$
cdf for crv: if $F(y)$ is continuous for $-\infty < y < \infty$
$P(Y = y) = 0$ and pf, pdf (probability density function) of crv is $f(y)$:

$$f(\cdot) = f(y) = \frac{dF(y)}{dy}F'(y), \text{ when derivative exists}$$

properties of pdf: (pdf doesn't have to be cont.)
1. $f(y) \geq 0$ for all y, $-\infty < y < \infty$
2. $\int_{-\infty}^{\infty} f(y)dy = 1$
pth quantile: if $0 < p < 1$, denoted by ϕ_p is the smallest $P(Y \leq \phi_q) = F(\phi_p) \geq p$. If Y is cont. ϕ_p is the smallest val such that $F(\phi_p = P(Y \leq \phi_p) = p)$
ϕ_p is also the 100pth percentile of Y.
$\phi_{.5}$ is median (50th percentile / quantile)

probability on interval:

$$P(a \leq Y \leq b) = \int_a^b f(y)dy$$

ev of crv if integral exists:

$$E(Y) = \int_{-\infty}^{\infty} yf(y)dy$$
$$E[g(Y)] = \int_{-\infty}^{\infty} g(y)f(y)dy$$

properties same: constant, closed under addition, and scalar multiplication

4.1 Uniform pd

If $\theta_1 < \theta_2$, crv Y has a continuous uniform pd on (θ_1, θ_2) iff:

$$f(y) = \begin{cases} \frac{1}{\theta_2 - \theta_1}, \theta_1 \leq y \leq \theta_2, \\ 0, \text{ elsewhere.} \end{cases}$$

mean and variance for $\theta_1 < \theta_2, (\theta_1, \theta_2)$

$$\mu = E(Y) = \frac{\theta_1 + \theta_2}{2} \text{ and } \sigma^2 = V(Y) = \frac{(\theta_2 - \theta_1)^2}{12}$$

4.2 Normal pd

4.3 Gamma pd

4.4 Beta pd

4.5 Moments and mgf

4.6 Tchebysheff

4.7 Mixed pd

5 Multivariate

6 Functions

7 Sampling Distributions

8 Estimation

9 Estimators

10 Hypothesis Testing

11 2.3

11.1