Lecture 18 - 10/3

Monday, October 03, 2022 9:57 AM

Week 1: ① Set theory. set operations, demorgan's law, Venn diagram, K-map

② Axioms of probability: ① Event axioms; ② Probability axioms

③ Counting size of sets: ① Principle of counting; ② Principle of multiplication; ③ Principle of over counting

④ Probability experiments with equally likely outcomes

If each member of Ω is equally likely, $P(A) = \frac{|A|}{|\Omega|}$, for any $A \subset \Omega$

Week 2: ① Random variables $X: \Omega \to R$

. Discrete RV: Support, realization, Probability mass function

② Mean, Variance, Standard deviation

. $\mu_X = E[X] = \sum_x x\, p_X(x)$

. $Var(X) = E[(X-\mu_X)^2] = E[X^2] - \mu_X^2$

. $\sigma_X = \sqrt{Var(X)}$

. LOTUS and implications: $E[aX^2 + bX + c] = aE[X^2] + bE[X] + c$

③ Conditional probability:

. $P(B|A) = \begin{cases} \frac{P(AB)}{P(A)}, & \text{if } P(A) > 0 \\ \text{undefined}, & \text{if } P(A) = 0 \end{cases}$

. It is a probability measure, i.e., satisfies axioms.

④ Independent events:

. A and B are independent if $P(AB) = P(A)P(B)$

. If A and B are independent, then $P(B|A) = P(B)$.

Week 3: ① Common pmfs:

1. Bernoulli distribution: $Ber(p)$, $p_X(1) = p$, $p_X(0) = 1-p$
2. Binomial distribution: $Bi(n,p)$, $p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}$ $k \in \{0,1,\dots,n\}$
3. Geometric distribution: $Geo(p)$, $p_X(k) = (1-p)^{k-1} p$ $k \in \{1,2,\dots\}$
4. Negative Binomial distribution: $NB(r,p)$, $p_X(n) = \binom{n-1}{r-1} p^r (1-p)^{n-r}$, $n \in \{0,1,2,\dots\}$
5. Poisson distribution: $Poi(\lambda)$, $p_X(k) = e^{-\lambda} \frac{\lambda^k}{k!}$, $k \in \{0,1,2,\dots\}$

4. Negative Binomial distribution: $NB(r,p)$, $p_X(n) = \binom{n-1}{r-1} p (1-p)$, $n \in \{0,1,2,\dots\}$

5. Poisson distribution: $Poi(\lambda)$, $p_X(k) = e^{-\lambda}\frac{\lambda^k}{k!}$, $k \in \{0,1,2,\dots\}$

③ Mean & variance of above distributions.

Week 4: ① Maximum likelihood estimator:

(i) distribution of random variable X belongs to a family of parametrized distributions

(ii) We are given an observation A, i.e., $X \in A$

(iii) We want to find the parameter that maximizes probability of A

② Inequalities:

Markov's inequality: $P(X \geq c) \leq \frac{E[X]}{c}$, $c > 0$ and X is a nonnegative random variable

Chebychev's inequality: $P(|X - E[X]| \geq a) \leq \frac{Var(X)}{a^2}$

③ Confidence interval:

. Suppose that each individual is in favor of policy A with probability p.

. We sample n individuals, X of them were in favor of policy A

. We do not know p.

(i) $\hat{p} = \frac{X}{n}$ is the point estimate of p

(ii) $P\left(p \in \left(\hat{p} - \frac{a}{2\sqrt{n}}, \hat{p} + \frac{a}{2\sqrt{n}}\right)\right) \geq 1 - \frac{1}{a^2}$

. We are $1 - \frac{1}{a^2}$ confident that p is within $\frac{a}{2\sqrt{n}}$ of $\hat{p}$.

④ law of total probability and Bayes' formula:

. $P(A) = \sum_{i=1}^{n} P(A \cap E_i) = \sum_{i=1}^{n} P(E_i) P(A|E_i)$, E_is partitions Ω, $P(E_i) > 0$

. $$P(E_i | A) = \frac{P(A|E_i)P(E_i)}{P(A|E_1)P(E_1) + \cdots + P(A|E_n)P(E_n)}$$

⑤ law of total expectation:

. $E[g(X)] = \sum_i P(E_i) E[g(X)|E_i]$ E_is partitions Ω, $P(E_i) > 0$

. $E[g(X)|A] = \sum_i g(u_i) P(X = u_i | A)$

. Example: $Var(A|X) = E[(X - E(X|A))^2 | A] = E[X^2|A] - (E[X|A])^2$

. If X is independent of A, then $E[g(X)|A] = E[g(X)]$

Week 5: ① Hypothesis testing

(i) The system generates random variable X

- If the system is in state H_0, the pmf of X is given by p_0
- If the system is in state H_1, the pmf of X is given by p_1

$p_0(k) = P(X=k|H_0)$ and $p_1(k) = P(X=k|H_1)$

(ii) We observe a realization of X, i.e, we observe $X=m$.

(iii) We guess the state of system, using our observation, based on the decision rule

② . ML maximizes likelihood of observation, MAP maximizes a posteriori probabilites

$$\text{ML: Given } X=k, \begin{cases} \text{declare } H_0 \text{ if } P(X=k|H_0) > P(X=k|H_1) \\ \text{declare } H_1 \text{ if } P(X=k|H_0) < P(X=k|H_1) \\ \text{declare either if } P(X=k|H_0) = P(X=k|H_1) \end{cases}$$

$$\text{MAP: Given } X=k \begin{cases} \text{declares } H_0 \text{ if } P(H_0|X=k) > P(H_1|X=k) \\ \text{declares } H_1 \text{ if } P(H_0|X=k) < P(H_1|X=k) \\ \text{declares eithe if } P(H_0|X=k) = P(H_1|X=k) \end{cases}$$

③ In ML, we compare likelihood ratio, $\Lambda(k) = \dfrac{p_1(k)}{p_0(k)} = \dfrac{P(X=k|H_1)}{P(X=k|H_0)}$ with 1.

$$\text{ML}: \Lambda(k): \begin{cases} \text{declare } H_0 \text{ if } \Lambda(k) < 1 \\ \text{declare } H_1 \text{ if } \Lambda(k) > 1 \\ \text{declare either if } \Lambda(k) = 1 \end{cases}$$

④ In MAP, we compare likelihood ratio, $\Lambda(k) = \dfrac{p_1(k)}{p_0(k)} = \dfrac{P(X=k|H_1)}{P(X=k|H_0)}$ with $\dfrac{\pi_0}{\pi_1}$.

$$\text{MAP}: \Lambda(k): \begin{cases} \text{declare } H_0 \text{ if } \Lambda(k) < \frac{\pi_0}{\pi_1} \\ \text{declare } H_1 \text{ if } \Lambda(k) > \frac{\pi_0}{\pi_1} \\ \text{declare either if } \Lambda(k) = \frac{\pi_0}{\pi_1} \end{cases}$$

⑤ $P_{\text{false alarm}}$: probability that decision rule decides H_1, given true state is H_0 $\rightsquigarrow$ conditional prob.

P_{miss}: probability that decision rule decides H_0, given true state is H_1 $\rightsquigarrow$ conditional prob.

P_{error}: probability that decision rule is not equal to the true state. $\rightsquigarrow$ regular probability.

error: probability that decision rule is not equal to the true state. regular probability.

1. **[20 points]** The two parts of this problem are unrelated. Spring 2015

   (a) Suppose that you have five distinct bowls, each can hold up to two oranges. If you have four indistinguishable oranges, how many ways are there to place the oranges in the bowls?

3. **[24 points]** A bag contains 4 orange and 6 blue t-shirts. Two t-shirts are chosen from the bag at random and moved into a second bag, which already contains one orange t-shirt. Then a t-shirt is chosen from the second bag at random. Fall 2014

   (a) (10 points) Let X denote the number of blue t-shirts moved from the first bag to the second bag. Find the pmf of X.

   (b) (8 points) What is the probability the t-shirt chosen from the second bag is blue?

   (c) (6 points) What is the conditional probability $X = 2$, given the t-shirt drawn from the second bag is blue?

3. [**6+8+10+10 points**] We have a bag of coins. Pick one coin from the bag and flip the same coin repeatedly. Fall 2018

   (a) The coin shows heads with probability p each time it is flipped. We flip the coin n times and denote the number of heads we observe by X. We will use $\hat{p} = \frac{X}{n}$ to estimate p. If we want to estimate p to within 0.1 with 99% confidence, how many times do we need to flip the coin?

   (b) The coin shows heads with probability p each time it is flipped. We observe the second head at the sixth trial. Compute the ML estimate of p. Show your work.

   (c) We flip the same coin three times and observe the total number of heads, X. Let H_0 be the hypothesis that the coin is fair. Let H_1 be the hypothesis that the coin is biased and shows heads with probability $\frac{1}{4}$. Write down the ML decision rule. Compute p_{miss}.

   (d) Suppose we know that 80% of the coins are fair and 20% of the coins show heads with probability $\frac{1}{3}$. If the first head shows at the second flip, what is the probability that we picked a fair coin?