UC Berkeley
Department of Electrical Engineering and Computer Sciences

EECS 126: Probability and Random Processes

Discussion 6
Fall 2021

1. Finite Boundary Times

Consider the random walk $S_n = \sum_{i=1}^{n} X_i$, where the X_i are iid with mean zero and variance 1 (note that they do not have to be discrete). Show that almost surely the random walk will leave the interval $[-a, a]$ in finite time.

Hint: Let T be the first time that the walk leaves the interval $[-a, a]$, and show that $\lim_{n\to\infty} P(T > n) = 0$.

Solution: We may apply CLT to see that $S_n/\sqrt{n} \xrightarrow{\text{d}} \mathcal{N}(0, 1)$. Hence,

$$P(T > n) \le P(|S_n/\sqrt{n}| \le a/\sqrt{n}) \to P(|\mathcal{N}(0, 1)| \le 0) = 0.$$

Since $P(T = \infty) = \lim_{n\to\infty} P(T > n)$, we conclude that $T < \infty$ almost surely.

2. Confidence Interval Comparisons

In order to estimate the probability of a head in a coin flip, p, you flip a coin n times, where n is a positive integer, and count the number of heads, S_n. You use the estimator $\hat{p} = S_n/n$.

(a) You choose the sample size n to have a guarantee

$$P(|\hat{p} - p| \ge \epsilon) \le \delta.$$

Using Chebyshev Inequality, determine n with the following parameters. Note that you should not have p in your final answer.

(i) Compare the value of n when $\epsilon = 0.05, \delta = 0.1$ to the value of n when $\epsilon = 0.1, \delta = 0.1$.

(ii) Compare the value of n when $\epsilon = 0.1, \delta = 0.05$ to the value of n when $\epsilon = 0.1, \delta = 0.1$.

(b) Now, we change the scenario slightly. You know that $p \in (0.4, 0.6)$ and would now like to determine the smallest n such that

$$P\Big(\frac{|\hat{p} - p|}{p} \le 0.05\Big) \ge 0.95.$$

Use the CLT to find the value of n that you should use. *Recall that the CLT states that the sum of IID random variables tends to a normal distribution with the sample mean and variance as it's parameters for n large enough.*

Solution:

(a) Chebyshev Inequality implies that:

$$P\Big(\Big|\frac{S_n}{n} - p\Big| \ge \epsilon\Big) \le \frac{\text{var}(S_n/n - p)}{\epsilon^2} = \frac{p(1-p)}{n\epsilon^2}$$

Thus, we set $\delta = p(1-p)/(n\epsilon^2)$ or $n = p(1-p)/(\delta\epsilon^2)$. Thus, when ϵ is reduced to half of its original value, n is changed to 4 times its original value, and when δ is reduced to half of its original value, n will be twice its original value. In order to be more concrete, we may maximize $p(1-p)/(\delta\epsilon^2)$ by letting $p = 1/2$. Thus, when $\epsilon = 0.1, \delta = 0.1$, $n = 250$. Letting $\delta = 0.05$ results in $n = 500$, while letting $\epsilon = 0.05$ results in $n = 1000$.

(b) Note that by the CLT:

$$\sqrt{n}\frac{\hat{p}-p}{\sqrt{p(1-p)}} \sim \mathcal{N}(0,1)$$

We are interested in the following:

$$\begin{aligned} P\Big(\frac{|\hat{p}-p|}{p} \le 0.05\Big) &\equiv P\Big(\Big|\sqrt{n}\frac{\hat{p}-p}{\sqrt{p(1-p)}}\Big| \le 0.05\frac{\sqrt{np}}{\sqrt{1-p}}\Big) \\ &\approx P\Big(|\mathcal{N}(0,1)| \le 0.05\frac{\sqrt{np}}{\sqrt{1-p}}\Big) \end{aligned}$$

Now, we use the condition that we want:

$$P\Big(|\mathcal{N}(0,1)| \le 0.05\frac{\sqrt{np}}{\sqrt{1-p}}\Big) \ge 0.95$$

This implies that $0.05\sqrt{np/(1-p)} \ge 2$ (note we use 2 here for simplicity, if you used 1.96, this is completely correct), or $n \ge 1600(1-p)/p$. We now use the fact that we know $p \in [0.4, 0.6]$. Since $p \in [0.4, 0.6]$, we can see that the value $(1-p)/p$ is maximized when $p = 0.4$. Thus, we note that $n \ge 1600(1-p)/p$ for all values of p, so the minimum value of n must be the maximum valid value of $1600(1-p)/p = 2400$.

3. **Characteristic Function Basics**

The definition of the characteristic function for random variable X is $\varphi_X(t) = \mathbb{E}[e^{itX}]$. It has many important properties - most notably that there is a bijection between the CDF (and therefore also PDF) of a random variable and its characteristic function. This problem goes over some of its basic properties.

(a) Let X be a Rademacher random variable, i.e. takes values ± 1 each with probability $1/2$. Show that $\varphi_X(t) = \cos(t)$.

(b) Let X be a uniform random variable on the interval $[a, b]$. Show that

$$\varphi_X(t) = \frac{e^{itb} - e^{ita}}{it(b-a)}.$$

. What happens if $b = -a$?

(c) Show that $\varphi_X(-t) = \overline{\varphi_X(t)}$, where the bar means take the complex conjugate. Use this fact to argue that if the distribution of X is symmetric about the origin, then the characteristic function is strictly real.

(d) Show that

$$\varphi_X^{(k)}(t)\Big|_{t=0} = i^k\mathbb{E}[X^k].$$

This can be particularly useful for computing higher moments of random variables.

(e) Show that that for independent $X_1, ..., X_n$ and scalars $a_1, ..., a_n$,

$$\varphi_{a_1X_1+...+a_nX_n}(t) = \varphi_{X_1}(a_1t) \cdot ... \cdot \varphi_{X_n}(a_nt).$$

This can be particularly useful for finding the distribution of $X+Y$ without having to deal with a convolution (in particular it tells us that convolution corresponds to multiplication in the fourier domain, a concept which may be familiar if you've taken some signals courses).

Solution:

(a) We have that

$$\varphi_X(t) = \frac{e^{it}+e^{-it}}{2} = \cos(t)$$

which follows from Euler's formula.

(b) We have

$$\begin{aligned}\varphi_X(t) = \mathbb{E}[e^{itX}] &= \int_a^b \frac{1}{b-a} e^{itx}dx \\ &= \left.\frac{e^{itx}}{it(b-a)}\right|_a^b \\ &= \frac{e^{itb}-e^{ita}}{it(b-a)}.\end{aligned}$$

If $b=-a$, then the above simplifies (using euler's formula again) to $\frac{\sin(tb)}{2tb}$.

(c) We have that $\varphi_X(-t) = \mathbb{E}[\cos(-tX)+i\sin(-tX)] = \mathbb{E}[\cos(tX)-i\sin(tX)] = \overline{\varphi_X(t)}$. For distributions that are symmetric about the origin, we know that $-X$ is distributed the same as X, so

$$\varphi_X(t) = \varphi_{-X}(t) = \varphi_X(-t) = \overline{\varphi_X(t)}.$$

The only way this can happen is if $\varphi_X(t)$ is a strictly real function.

(d) Using the fourier expansion of e^x and linearity of expectation, one has $\varphi_X(t) = 1 + \frac{it\mathbb{E}[X]}{1} + \frac{(it)^2\mathbb{E}[X^2]}{2!} + ...$ Taking the derivative k times and setting $t=0$ yields the result.

(e) We have that

$$\begin{aligned}\varphi_{a_1X_1+...+a_nX_n}(t) &= \mathbb{E}[e^{it(a_1X_1+...+a_nX_n)}] \\ &= \mathbb{E}[e^{ita_1X_1}e^{ita_2X_2}\cdots e^{ita_nX_n}] \\ &= \mathbb{E}[e^{ita_1X_1}]\cdots\mathbb{E}[e^{ita_nX_n}] = \varphi_{X_1}(a_1t)\cdots\varphi_{X_n}(a_nt).\end{aligned}$$