UC Berkeley
Department of Electrical Engineering and Computer Sciences

EECS 126: Probability and Random Processes

Homework 5
Spring 2024

1. **Polar coordinates**

Let X and Y be independent standard normal random variables. The pair (X, Y) can be described in polar coordinates in terms of random variables $R \geq 0$ and $\Theta \in [0, 2\pi]$, so that

$$X = R\cos\Theta, \quad Y = R\sin\Theta.$$

(a) Show that the joint CDF of R, Θ is given by $\frac{1}{2\pi}\int_0^\theta \int_0^r e^{-r'^2/2} r' dr' d\theta'$.

(b) Differentiate the joint CDF to find the pdf of R, Θ. Find immediately the marginal distribution of Θ.

(c) Show that R^2 has an exponential distribution with parameter $1/2$.

Solution:

The joint PDF of X and Y is

$$f_{X,Y}(x, y) = f_X(x) f_Y(y) = \frac{1}{2\pi} e^{-(x^2+y^2)/2}.$$

We first find the joint CDF of R and Θ. Fix some $r > 0$ and some $\theta \in [0, 2\pi]$, and let A be the set of points (x, y) whose polar coordinates (r, θ) satisfy $0 \leq r' \leq r$ and $0 \leq \theta' \leq \theta$; note that the set A is a sector of a circle of radius r with angle θ. We have

$$F_{R,\Theta}(r, \theta) = P(R \leq r', \Theta \leq \theta') = P((X, Y) \in A)$$
$$= \frac{1}{2\pi} \int\int_{(x,y)\in A} e^{-(x^2+y^2)/2} dx dy = \frac{1}{2\pi}\int_0^\theta \int_0^r e^{-r'^2/2} r' dr' d\theta',$$

where the last equality is obtained by transforming to polar coordinates. We then differentiate, to find that

$$f_{R,\Theta}(r, \theta) = \frac{\partial^2 F_{R,\Theta}(r, \theta)}{\partial r \partial \theta} = \frac{r}{2\pi} e^{-r^2/2}, \quad r \geq 0, \theta \in [0, 2\pi].$$

Thus,

$$f_R(r) = \int_0^{2\pi} f_{R,\Theta}(r, \theta) d\theta = r e^{-r^2/2}, \quad r \geq 0.$$

Furthermore,

$$f_{\Theta|R}(\theta|r) = \frac{f_{R,\Theta}(r, \theta)}{f_R(r)} = \frac{1}{2\pi}, \quad \theta \in [0, 2\pi].$$

Since the conditional PDF $f_{\Theta|R}$ of Θ is unaffected by the value of the conditioning variable R. It follows that it is also equal to the unconditional PDF f_Θ. In particular, $f_{R,\Theta}(r,\theta) = f_R(r) f_\Theta(\theta)$, so that R and Θ are independent.

Let $t \geq 0$. We have

$$P(R^2 \geq t) = P(R \geq \sqrt{t}) = \int_{\sqrt{t}}^{\infty} r e^{-r^2/2} dr = \int_{t/2}^{\infty} e^{-u} du = e^{-t/2},$$

where we have used the change of variables $u = r^2/2$. By differentiating, we obtain

$$f_{R^2}(t) = \frac{1}{2} e^{-t/2}, \quad t \geq 0.$$

2. **Gaussian practice**

   Let (η, ζ) be joint Gaussian random variables, such that $\mathbb{E}[\xi] = \mathbb{E}[\eta] = 0$, $\text{Var}(\xi) = \text{Var}(\eta) = 1$, and $\text{Cov}(\eta, \xi) = \rho$.

   (a) Are the random variables $\eta - \xi$ and $\eta + \xi$ independent?

   (b) Show that $\mathbb{E}[\min(\xi, \eta)] = -\sqrt{\frac{1-\rho}{\pi}}$.

   Hint: $\int x e^{-\alpha x^2} dx = -\frac{e^{-\alpha x^2}}{2\alpha} + c$.

   Solution:

   (a) They are independent.

   By property of Gaussian, $\begin{pmatrix} \eta - \xi \\ \eta + \xi \end{pmatrix}$ is still joint Gaussian and its covariance matrix is given by

$$\begin{aligned} &\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}^\top \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \\ = &\begin{pmatrix} 2(1+\rho) & 0 \\ 0 & 2(1-\rho) \end{pmatrix}. \end{aligned}$$

   It follows from property of Gaussian that $\eta - \xi$ and $\eta + \xi$ are independent.

   (b) Notice that $\min(\xi, \eta) = \frac{\xi + \eta - |\xi - \eta|}{2}$. It follows that

$$\mathbb{E}[\min(\xi, \eta)] = \mathbb{E}[(\xi + \eta - |\xi - \eta|)/2] = -\mathbb{E}[|\xi - \eta|]/2.$$

   For $\mathbb{E}[|\xi - \eta|]$, we calculate the integral

$$\begin{aligned} \mathbb{E}[|\xi - \eta|] &= \frac{1}{\sqrt{2\pi}\sqrt{2(1-\rho)}} \int_{-\infty}^{+\infty} |x| e^{-\frac{x^2}{4(1-\rho)}} dx \\ &= \frac{2}{\sqrt{2\pi}\sqrt{2(1-\rho)}} \int_{0}^{+\infty} x e^{-\frac{x^2}{4(1-\rho)}} dx \\ &= -\frac{1}{\sqrt{\pi(1-\rho)}} \cdot \frac{4(1-\rho)}{2} \cdot e^{-\frac{x^2}{4(1-\rho)}} \Bigg|_0^\infty \\ &= 2\sqrt{\frac{1-\rho}{\pi}}. \end{aligned}$$

3. **Matrix Sketching**

Matrix sketching is an important technique in randomized linear algebra for doing large computations efficiently. For example, to compute $\mathbf{A}^T \times \mathbf{B}$ for two large matrices $\mathbf{A}$ and $\mathbf{B}$, we can use a random sketch matrix $\mathbf{S}$ to compute a "sketch" $\mathbf{SA}$ of $\mathbf{A}$, and a sketch $\mathbf{SB}$ of $\mathbf{B}$. Such a sketching matrix has the property that

$$\mathbf{S}^T\mathbf{S} \approx \mathbf{I},$$

so that the approximate multiplication $(\mathbf{SA})^T(\mathbf{SB}) = \mathbf{A}^T\mathbf{S}^T\mathbf{SB}$ is close to $\mathbf{A}^T\mathbf{B}$.

In this problem, we will discuss two popular sketching schemes and understand how they help in approximate computation. Let $\hat{\mathbf{I}} = \mathbf{S}^T\mathbf{S}$, and let the dimension of the sketch matrix $\mathbf{S}$ be $d \times n$ (where typically $d \ll n$).

a. **Gaussian sketch**. Let the sketch matrix be

$$\mathbf{S} = \frac{1}{\sqrt{d}} \begin{bmatrix} S_{1,1} & \cdots & S_{1,n} \\ \vdots & \ddots & \vdots \\ S_{d,1} & \cdots & S_{d,n} \end{bmatrix},$$

where the $S_{i,j}$ are chosen i.i.d. from $\mathcal{N}(0,1)$ for all $i \in [1,d]$ and $j \in [1,n]$. Show that the elementwise mean and variance of the matrix $\hat{\mathbf{I}}$, as functions of d, are

$$\mathbb{E}(\hat{I}_{i,j}) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

$$\text{var}(\hat{I}_{i,j}) = \begin{cases} \frac{2}{d} & \text{if } i = j \\ \frac{1}{d} & \text{otherwise.} \end{cases}$$

You can use without proof the fact that $\mathbb{E}(Z^4) = 3$ for $Z \sim \mathcal{N}(0,1)$.

b. **Count sketch**. For each column $j \in [1,n]$ of $\mathbf{S}$, choose a row i uniformly randomly from $[1,d]$. Set

$$S_{i,j} = \begin{cases} 1 & \text{with probability } \frac{1}{2} \\ -1 & \text{with probability } \frac{1}{2}, \end{cases}$$

and assign $S_{k,j} = 0$ for all $k \neq i$. An example of a 3×8 count sketch matrix is

$$\begin{bmatrix} 0 & -1 & 1 & 0 & 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 & -1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 \end{bmatrix}.$$

Show that the elementwise mean and variance of the matrix $\hat{\mathbf{I}}$ are

$$\mathbb{E}(\hat{I}_{i,j}) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

$$\text{var}(\hat{I}_{i,j}) = \begin{cases} 0 & \text{if } i = j \\ \frac{1}{d} & \text{otherwise.} \end{cases}$$

Note that for sufficiently large d, the matrix $\hat{\mathbf{I}}$ is close to the identity matrix in both cases. We use this fact in the lab to do an approximate matrix multiplication.

Solution:

a. For the Gaussian sketch matrix $\mathbf{S}$, we have

$$\hat{I}_{i,j} = \frac{1}{d} \sum_{k=1}^{d} S_{k,i} S_{k,j}.$$

By the linearity of expectation, and the $S_{k,i}$ being drawn i.i.d. from $\mathcal{N}(0,1)$, we get

$$\mathbb{E}(\hat{I}_{i,j}) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases}$$

Then, by the definition of variance, we have

$$\begin{aligned} d^2 \operatorname{var}(\hat{I}_{i,j}) &= \mathbb{E}[(d\hat{I}_{i,j})^2] - \mathbb{E}[d\hat{I}_{i,j}]^2 \\ &= \mathbb{E}\left[\left(\sum_{k=1}^{d} S_{k,i} S_{k,j}\right)^2\right] - d^2\, \mathbb{1}_{i=j}\,. \end{aligned}$$

Now we consider the two cases of $i = j$ and $i \neq j$, starting with the former:

$$\begin{aligned} d^2 \operatorname{var}(\hat{I}_{i,i}) &= \sum_{k=1}^{d} \mathbb{E}(S_{k,i}^4) + \sum_{k \neq \ell} \mathbb{E}(S_{k,i}^2)\, \mathbb{E}(S_{\ell,i}^2) - d^2 \\ &= 3d + d(d-1) - d^2 = 2d. \end{aligned}$$

For the case of $i \neq j$, we can use the independence of $S_{k,i}$ and $S_{k,j}$:

$$\begin{aligned} d^2 \operatorname{var}(\hat{I}_{i,j}) &= \sum_{k=1}^{d} \mathbb{E}(S_{k,i}^2)\, \mathbb{E}(S_{k,j}^2) + \sum_{k \neq \ell} \mathbb{E}(S_{k,i})\, \mathbb{E}(S_{k,j})\, \mathbb{E}(S_{\ell,i})\, \mathbb{E}(S_{\ell,j}) \\ &= d + 0 = d. \end{aligned}$$

Thus the elementwise variance is

$$\operatorname{var}(\hat{I}_{i,j}) = \begin{cases} \frac{2}{d} & \text{if } i = j \\ \frac{1}{d} & \text{otherwise.} \end{cases}$$

b. For the count sketch matrix $\mathbf{S}$, we have

$$\hat{I}_{i,j} = \sum_{k=1}^{d} S_{k,i} S_{k,j}.$$

By construction of $\mathbf{S}$, the diagonal terms $\hat{I}_{i,i}$ are always 1, so their mean is 1 and their variance is 0, and we only need to worry about the non-diagonal terms.

We also note that in $\mathbf{S}$, entries in a row are independent, but entries in a column are dependent. (There can only be one nonzero entry in one column.) Moreover, for all $i \neq j$,

$$S_{k,i} S_{k,j} = \begin{cases} 1 & \text{with probability } \frac{1}{2d^2} \\ -1 & \text{with probability } \frac{1}{2d^2} \\ 0 & \text{with probability } 1 - \frac{1}{d^2}. \end{cases}$$

Thus the elementwise expectation is

$$\mathbb{E}(\hat{I}_{i,j}) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases}$$

Now, for $i \neq j$, using the fact that $\mathbb{E}[\hat{I}_{i,j}]^2 = 0$,

$$\begin{aligned} \operatorname{var}(\hat{I}_{i,j}) &= \mathbb{E}\left[\left(\sum_{k=1}^{d} S_{k,i} S_{k,j}\right)^2\right] \\ &= \sum_{k=1}^{d} \mathbb{E}(S_{k,i}^2)\, \mathbb{E}(S_{k,j}^2) + \sum_{k \neq \ell} \mathbb{E}(S_{k,i} S_{\ell,i})\, \mathbb{E}(S_{k,j} S_{\ell,j}) \\ &= \sum_{k=1}^{d} \frac{1}{d^2} + 0 = \frac{1}{d}. \end{aligned}$$

The term 0 in the last step comes from the fact that in any column j, the product of two elements $S_{k,j} S_{\ell,j} = 0$, since only one can be nonzero. Thus the elementwise variance is

$$\operatorname{var}(\hat{I}_{i,j}) = \begin{cases} 0 & \text{if } i = j \\ \frac{1}{d} & \text{otherwise.} \end{cases}$$

4. **Coupon Collector Bounds**

Recall the coupon collector's problem, in which there are n different types of coupons. Every box contains a single coupon, and we let the random variable X be the number of boxes bought until one of every type of coupon is obtained. The expected value of X is nH_n, where $H_n := \sum_{i=1}^{n} \frac{1}{i}$ is the *harmonic number* of order n, which satisfies the inequality

$$\ln n \leq H_n \leq \ln n + 1.$$

a. Use Markov's inequality in order to show that

$$\mathbb{P}(X > 2nH_n) \leq \frac{1}{2}.$$

b. Use Chebyshev's inequality in order to show that

$$\mathbb{P}(X > 2nH_n) \leq \frac{\pi^2}{6(\ln n)^2}.$$

Note: You can use Euler's solution to the Basel problem, the identity $\sum_{i=1}^{\infty} \frac{1}{i^2} = \frac{\pi^2}{6}$.

c. Define appropriate events and use the union bound in order to show that

$$\mathbb{P}(X > 2nH_n) \leq \frac{1}{n}.$$

Note: $a_n = (1 - \frac{1}{n})^n$ is a strictly increasing sequence with limit e^{-1}.

Solution:

a. We are given $\mathbb{E}(X) = nH_n$, so

$$\mathbb{P}(X > 2nH_n) \leq \frac{\mathbb{E}(X)}{2nH_n} = \frac{1}{2}.$$

b. We can write X as an independent sum $\sum_{i=1}^{n} X_i$, where $X_i \sim \text{Geometric}(\frac{n-i+1}{n})$, so

$$\text{var}(X) = \sum_{i=1}^{n} \text{var}(X_i) < \sum_{i=1}^{n} \left(\frac{n}{n-i+1}\right)^2 = \sum_{i=1}^{n} \left(\frac{n}{i}\right)^2 < \frac{\pi^2 n^2}{6}.$$

Using Chebyshev's inequality, we have that

$$\mathbb{P}(X > 2nH_n) \leq \mathbb{P}(|X - nH_n| > nH_n) \leq \frac{\text{var}(X)}{(nH_n)^2} < \frac{\pi^2}{6H_n^2} \leq \frac{\pi^2}{6(\ln n)^2}.$$

c. Let A_i be the event that we fail to get box i after $2nH_n$ tries.

$$\mathbb{P}(A_i) \leq \left(\frac{n-1}{n}\right)^{2nH_n} = \left[\left(1 - \frac{1}{n}\right)^n\right]^{2H_n} < e^{-2H_n} \leq e^{-2\ln n} = \frac{1}{n^2}.$$

Now, by the union bound, we can conclude that

$$\mathbb{P}(X > 2nH_n) = \mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} \mathbb{P}(A_i) \leq \frac{1}{n}.$$

5. **Subgaussian Random Variables**

We say that a random variable Y is *subgaussian* with parameter σ^2 (written $Y \in SG(\sigma^2)$) if

$$\mathbb{E}\, e^{\lambda Y} \leq \exp\left(\frac{1}{2}\lambda^2\sigma^2\right), \quad \text{for all } \lambda \in \mathbb{R}.$$

Subgaussian random variables consist a whole class of random variables which have a similar tail with Gaussian random variables, and thus they share similar concentration properties. We are going to examine some properties of subgaussian random variables in this sequel of problems.

Prove the following:

a. If $Y \sim N(0, \sigma^2)$, then $Y \in SG(\sigma^2)$.

b. If $\Pr(Y = 1) = \Pr(Y = -1) = 1/2$, then $Y \in SG(1)$.
Hint: use Taylor expansion for the exponentials and observe what is happening to the odd terms.

c. If $Y \in SG(\sigma^2)$, then $\mathbb{E}[Y] = 0$.
Hint: what is the power series for $\mathbb{E}\, e^{\lambda Y}$? How can you extract $\mathbb{E}\, Y$ from the power series?

d. If $Y \in SG(\sigma^2)$, then $\mathrm{var}(Y) \leq \sigma^2$.

e. If $Y_1, \ldots, Y_n$ are independent random variables with $Y_i \in SG(\sigma_i^2)$ for $i = 1, \ldots, n$, and $a_1, \ldots, a_n$ are scalars then

$$a_1Y_1 + \ldots + a_nY_n \in SG(a_1^2\sigma_1^2 + \ldots + a_n^2\sigma_n^2).$$

f. If $Y \in SG(\sigma^2)$, then

$$\Pr(Y \geq t) \leq \exp\left(-\frac{t^2}{2\sigma^2}\right) \quad \text{and} \quad \Pr(-Y \geq t) \leq \exp\left(-\frac{t^2}{2\sigma^2}\right)$$

for every $t \geq 0$.

Solution:

a. The moment generating function of $Y \sim N(0, \sigma^2)$ is

$$M_Y(\lambda) = \mathbb{E}\, e^{\lambda Y} = \exp\left(\frac{1}{2}\lambda^2\sigma^2\right).$$

So clearly $Y \in SG(\sigma^2)$.

b. In this case, $\sigma^2 = 1$ and

$$\begin{aligned}\mathbb{E}\, e^{\lambda Y} &= \frac{e^{-\lambda} + e^{\lambda}}{2} \\ &\leq \exp\left(\frac{1}{2}\lambda^2\right) \\ &= \exp\left(\frac{1}{2}\lambda^2\sigma^2\right),\end{aligned}$$

where the inequality is true by using Taylor expansions and the simple inequality $2^k k! \leq (2k)!$, as follows

$$\begin{aligned}\frac{e^{-\lambda}+e^{\lambda}}{2} &= \frac{1}{2}\sum_{i=0}^{\infty}(1+(-1)^i)\frac{\lambda^i}{i!} \\ &= \sum_{k=0}^{\infty}\frac{\lambda^{2k}}{(2k)!} \\ &\leq \sum_{k=0}^{\infty}\frac{(\frac{1}{2}\lambda^2)^k}{k!} \\ &= \exp\left(\frac{1}{2}\lambda^2\right).\end{aligned}$$

c. Note that $\mathbb{E}\,e^{\lambda Y} = 1 + \lambda\,\mathbb{E}\,Y + \frac{\lambda^2\,\mathbb{E}\,Y^2}{2} + \ldots$. Therefore, as $\lambda \to 0^+$ we have that

$$\frac{\mathbb{E}\,e^{\lambda Y}-1}{\lambda} \to \mathbb{E}\,Y \leq \frac{\exp\left(\frac{1}{2}\lambda^2\sigma^2\right)-1}{\lambda} \to 0,$$

and similarly, as $\lambda \to 0^-$ we have that

$$\frac{\mathbb{E}\,e^{\lambda Y}-1}{\lambda} \to \mathbb{E}\,Y \geq \frac{\exp\left(\frac{1}{2}\lambda^2\sigma^2\right)-1}{\lambda} \to 0,$$

hence we can conclude that $\mathbb{E}\,Y = 0$.

d. As $\lambda \to 0$ we have that

$$\frac{\mathbb{E}\,e^{\lambda Y}-1}{\lambda^2} \to \mathbb{E}\,Y^2 \leq \frac{\exp\left(\frac{1}{2}\lambda^2\sigma^2\right)-1}{\lambda^2} \to \sigma^2,$$

and because $\mathbb{E}\,Y = 0$, we conclude that $\mathrm{var}(Y) = \mathbb{E}\,Y^2 \leq \sigma^2$.

e. Because of the independence assumption the moment generating function of the sum, can be written as the product of the moment generating functions of each summand, and therefore we have that

$$\begin{aligned}M_{a_1Y_1+\ldots+a_nY_n}(\lambda) &= M_{a_1Y_1}(\lambda)\ldots M_{a_nY_n}(\lambda) \\ &= M_{Y_1}(a_1\lambda)\ldots M_{Y_n}(a_n\lambda) \\ &\leq \exp\left(\frac{1}{2}\lambda^2a_1^2\sigma_1^2\right)\ldots\exp\left(\frac{1}{2}\lambda^2a_n^2\sigma_n^2\right) \\ &= \exp\left(\frac{1}{2}\lambda^2(a_1^2\sigma_1^2+\ldots+a_n^2\sigma_n^2)\right).\end{aligned}$$

f. For any $\lambda > 0$ we have that,

$$\begin{aligned}\Pr(Y \geq t) &= \Pr(\lambda Y \geq \lambda t) \\ &= \Pr(e^{\lambda Y} \geq e^{\lambda t}) \\ &\leq \frac{\mathbb{E}\,e^{\lambda Y}}{e^{\lambda t}} \\ &\leq \exp\left(\frac{1}{2}\lambda^2\sigma^2 - \lambda t\right).\end{aligned}$$

Optimizing this upper bound with respect to $\lambda > 0$, yields that the best λ is $\lambda^* = t/\sigma^2$, from which the result follows.

In order to upper bound $\Pr(-Y \geq t)$, just observe that $Y \in SG(\sigma^2)$ implies that $-Y \in SG(\sigma^2)$, and hence the upper bound follows directly from the work that we have already done.

6. **The Weak Law of Large Numbers**

Let $X_1, X_2, \ldots$ be a sequence of i.i.d. random variables with common mean μ and MGF M_X. We assume that $M_X(s)$ is finite when $s \in (-d, d)$ for some $d > 0$. Let

$$\bar{X}_n := \frac{X_1 + \ldots + X_n}{n}.$$

a. Show that the transform (or MGF) associated with $\bar{X}_n$ satisfies

$$M_{\bar{X}_n}(s) = M_X(s/n)^n.$$

b. Suppose that the transform $M_X(s)$ has a first-order Taylor series expansion around $s = 0$ of the form

$$M_X(s) = a + bs + o(s),$$

where $o(s)$ is a function that satisfies $\lim_{s\to 0} o(s)/s = 0$. Find a and b in terms of μ.

c. Show that for all $s \in (-d, d)$,

$$\lim_{n\to\infty} M_{\bar{X}_n}(s) = e^{\mu s}.$$

Hint: If $(a_n)_{n\in\mathbb{N}}$ is a sequence of real numbers converging to a, then $\lim_{n\to\infty}(1+\frac{a_n}{n})^n = e^a$.

d. Deduce that $\bar{X}_n \xrightarrow{d} \mu$. Note that the pointwise convergence of MGFs is equivalent to convergence in distribution.

Solution:

a. Let $S_n = X_1 + \cdots + X_n$. Then, as the sequence of random variables is i.i.d., we have that

$$M_{S_n}(s) = M_X(s)^n.$$

In addition,

$$M_{\bar{X}_n}(s) = M_{S_n/n}(s) = \mathbb{E}\left(\exp\left(s\frac{S_n}{n}\right)\right) = M_{S_n}(s/n) = M_X(s/n)^n.$$

b. $a = 1$ and $b = \mu$.

c. Taking the hint given,

$$\lim_{n\to\infty} M_{\bar{X}_n}(s) = \lim_{n\to\infty}\left(1 + \frac{\mu s + s\frac{o(1/n)}{1/n}}{n}\right)^n = e^{\mu s}.$$

d. Let $Y = \mu$ be a constant random variable, and observe that $M_Y(s) = e^{s\mu}$. Since $M_{\bar{X}_n} \to M_Y$ as $n \to \infty$, as the MGF uniquely determines the distribution, we can deduce that $\bar{X}_n \xrightarrow{d} Y = \mu$ as $n \to \infty$.