High Dimensional Analysis: Random Matrices and Machine Learning Problems
(Work in progress).

Solutions by Gustavo Fuentes

Contents

1 Sheet 1 solutions

Solution 1.1. Define

$$I = \int_{\mathbb{R}} \exp(-t^2)dt$$

Then:

$$I^2 = \left(\int_{\mathbb{R}} \exp(-t^2)dt\right)^2 = \left(\int_{\mathbb{R}} \exp(-t^2)dt\right)\left(\int_{\mathbb{R}} \exp(-t^2)dt\right)$$

t is a dummy variable inside the integral

$$= \left(\int_{\mathbb{R}} \exp(-t^2)dt\right)\left(\int_{\mathbb{R}} \exp(-s^2)ds\right) = \int_{\mathbb{R}} \exp(-t^2)\exp(-s^2)dtds = \int_{\mathbb{R}}\int_{\mathbb{R}} \exp(-(t^2+s^2))dtds$$

Using the following change of variables

$$\begin{aligned}
&t = r\cos\theta \\
&s = r\sin\theta \\
&\Rightarrow t^2 + s^2 = (r\cos\theta)^2 + (r\sin\theta)^2 = r^2\left(\cos^2\theta + \sin^2\theta\right) = r^2 \\
&t^2 + s^2 = r^2
\end{aligned}$$

Where $r \in [0, \infty)$ and $\theta \in [0, 2\pi)$, the Jacobian is given by:

$$\begin{aligned}
&|\mathcal{J}| = \det\begin{bmatrix} \frac{\partial t}{\partial r} & \frac{\partial t}{\partial \theta} \\ \frac{\partial s}{\partial r} & \frac{\partial t}{\partial \theta} \end{bmatrix} = \det\begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix} = r\cos^2\theta + r\sin^2\theta = r \\
&|\mathcal{J}| = r
\end{aligned}$$

$$I^2 = \int_0^{2\pi}\int_0^{\infty} r\exp -r^2 dr d\theta = 2\pi\int_0^{\infty} r\exp -r^2 dr$$

Using the following change of variables, $u = r^2 \Rightarrow du = -2rdr$

$$= 2\pi\int_0^{\infty} \exp -u\frac{-1}{2}du = -\pi\left(\exp\left(-\infty\right) + \exp\left(0\right)\right) = \pi$$

then

$$\begin{aligned}
&I^2 = \pi \\
&I = \sqrt{\pi}
\end{aligned}$$

Hence:

$$\int_{\mathbb{R}} \exp(-t^2)dt = \sqrt{\pi}$$

Solution 1.2. (a)

$$
\begin{aligned}
E[x^0] &= \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right) dt = \frac{1}{\sqrt{2\pi\sigma^2}}\sqrt{2\pi\sigma^2} = 1 \\
E[x^0] &= 1
\end{aligned}
$$

The pdf ψ is normalized.

$$
\begin{aligned}
E[x^1] &= \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi\sigma^2}} t \exp\left(-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right) dt = \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \left[\int_{\mathbb{R}} \left(\frac{t-\mu}{\sigma}\right) \exp\left(-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right) dt + \mu \int_{\mathbb{R}} \exp\left(-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right) dt\right] = \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \mu\sqrt{2\pi\sigma^2} = \mu
\end{aligned}
$$

Then μ is the first moment or the mean of the pdf ψ.

(b) Using the Stein's identity:

$$
\begin{aligned}
E[g(x)(x-\mu)] &= \sigma^2 E\left[\frac{\mathrm{d}g(x)}{\mathrm{d}x}\right] \\
\Rightarrow E[g(x)x] &= \mu E[g(x)] + \sigma^2 E\left[\frac{\mathrm{d}g(x)}{\mathrm{d}x}\right]
\end{aligned}
$$

For $g(x) = x^{n-1} \Rightarrow \frac{\mathrm{d}g(x)}{\mathrm{d}x} = (n-1)x^{n-2}$

$$
\begin{aligned}
&\Rightarrow E[x^{n-1} \cdot x] = \mu E[x^{n-1}] + \sigma^2(n-1)E[x^{n-2}] \\
&\therefore E[x^n] = \mu E[x^{n-1}] + \sigma^2(n-1)E[x^{n-2}]
\end{aligned}
$$

(c)

$$
\begin{aligned}
&E[x^2] = \mu E[x^1] + \sigma^2 E[x^0] \\
&\therefore E[x^2] = \mu^2 + \sigma^2 \\
&E[x^3] = \mu E[x^2] + (3-1)\sigma^2 E[x^1] = \mu(\mu^2+\sigma^2) + 2\sigma^2\mu \\
&\therefore E[x^3] = \mu^3 + 3\mu\sigma^2 \\
&E[x^4] = \mu E[x^3] + (4-1)\sigma^2 E[x^2] = \mu(\mu^3 + 3\mu\sigma^2) + 3\sigma^2(\mu^2+\sigma^2) \\
&\therefore E[x^4] = \mu^4 + 6\mu^2\sigma^2 + 3\sigma^4
\end{aligned}
$$

(d) For $\mu = 0$

$$
\begin{aligned}
E[x^1] &= 0 \\
E[x^2] &= \sigma^2 \\
E[x^3] &= 0 \\
E[x^4] &= 3\sigma^4
\end{aligned}
$$

In general

$$E[x^n] = (n-1)\sigma^2 E[x^{n-2}] = (n-1)\sigma^2(n-3)\sigma^2 E[x^{n-4}] =$$
$$= (n-1)\sigma^2(n-3)\sigma^2(n-5)\sigma^2 E[x^{n-6}] = \cdots = (n-1)(n-3)\ldots 5\cdot 3\cdot 1\sigma^{2n}$$

if n is odd then $E[x^n] = 0$

$$\therefore E[x^n] = \begin{cases} 0 & \text{if n is odd} \\ \sigma^n(n-1)!! & \text{if n is even} \end{cases}$$

(e)

$$E[|x|] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} |t| \exp\left(-\frac{1}{2}t^2\right)dt = \frac{2}{\sqrt{2\pi}}\int_0^{\infty} t\exp\left(-\frac{1}{2}t^2\right)dt =$$
$$= \sqrt{\frac{2}{\pi}}\left(-\exp\left(-\frac{t^2}{2}\right)\right)\Big|_0^{\infty} = \sqrt{\frac{2}{\pi}}(1-0)$$
$$\therefore E[|x|] = \sqrt{\frac{2}{\pi}}$$

Solution 1.3. (a) Consider the following change of variables $t = y + s \Rightarrow dt = ds$. And the new integration limits are: $y = y + s \Rightarrow s = 0$ and $\infty = y + s \Rightarrow s = \infty$, hence:

$$\int_y^{\infty} \exp -t^2 dt = \int_0^{\infty} \exp-(y+s)^2 dt =$$
$$= \int_0^{\infty} \exp\left(-s^2 - 2ys - y^2\right)dt = \exp\left(-y^2\right)\int_0^{\infty} \exp\left(-s^2\right)\exp\left(-2ys\right)dt =$$
$$\leq \exp\left(-y^2\right)\int_0^{\infty} \exp\left(-s^2\right)dt$$

Notice that, by symmetry:

$$\int_{-\infty}^{\infty} \exp\left(-t^2\right)dt = 2\int_0^{\infty} \exp\left(-t^2\right)dt$$
$$\Rightarrow \int_0^{\infty} \exp\left(-t^2\right)dt = \frac{\sqrt{\pi}}{2}$$

Then:

$$\exp\left(-y^2\right)\int_0^{\infty} \exp\left(-s^2\right)dt = \frac{\sqrt{\pi}}{2}\exp\left(-y^2\right)$$
$$\therefore \int_y^{\infty} \exp -t^2 dt \leq \frac{\sqrt{\pi}}{2}\exp\left(-y^2\right)$$

(b) Notice that if $p \geq 3 \Rightarrow 1 > \frac{1}{\sqrt{p-1}}$ and $\frac{p-1}{2} \geq 1$. On the other hand $0 \leq t \leq 1 \Rightarrow 0 \leq t^2 \leq 1 \Rightarrow -1 \leq -t^2 \leq 0$. The integrand is positive, then if we reduce the range of integration:

$$\int_0^1 \left(1-t^2\right)^{\frac{p-1}{2}} dt \geq \int_0^{\frac{1}{\sqrt{p-1}}} \left(1-t^2\right)^{\frac{p-1}{2}} dt$$

Using the Bernoulli's inequality for $b = \frac{p-1}{2}$ and $a = -t^2$, then

$$\int_0^{\frac{1}{\sqrt{p-1}}} \left(1-t^2\right)^{\frac{p-1}{2}} dt \geq \int_0^{\frac{1}{\sqrt{p-1}}} \left(1 - t^2 \frac{p-1}{2}\right) dt =$$

$$= \int_0^{\frac{1}{\sqrt{p-1}}} dt - \frac{p-1}{2} \int_0^{\frac{1}{\sqrt{p-1}}} t^2 dt = \frac{1}{\sqrt{p-1}} - \frac{p-1}{2}\frac{1}{3}\frac{1}{(p-1)^{3/2}} =$$

$$\frac{1}{\sqrt{p-1}} - \frac{p-1}{2}\frac{1}{6}\frac{1}{(p-1)^{1/2}} = \frac{5}{6}\frac{1}{\sqrt{p-1}} > \frac{1}{2\sqrt{p-1}}$$

$$\therefore \int_0^1 \left(1-t^2\right)^{\frac{p-1}{2}} dt \geq \int_0^{\frac{1}{\sqrt{p-1}}} \left(1-t^2\right)^{\frac{p-1}{2}} dt > \frac{1}{2\sqrt{p-1}}$$

(c) The volume ratio is

$$\frac{\text{vol}[B_{p-1}]}{\text{vol}[B_p]} = \frac{\frac{\pi^{\frac{p-1}{2}}}{\Gamma\left(\frac{p-1}{2}+1\right)}}{\frac{\pi^{\frac{p}{2}}}{\Gamma\left(\frac{p}{2}+1\right)}} = \frac{1}{\sqrt{\pi}} \frac{\Gamma\left(\frac{p}{2}+1\right)}{\Gamma\left(\frac{p-1}{2}+1\right)} =$$

$$\approx \frac{1}{\sqrt{\pi}} \frac{\sqrt{2\pi\frac{p}{2}}}{\sqrt{2\pi\frac{p-1}{2}}} \frac{\left(\frac{p}{2e}\right)^{p/2}}{\left(\frac{p}{2e}\right)^{(p-1)/2}} = \frac{1}{\sqrt{\pi}}\sqrt{\frac{p}{p-1}} \left(\frac{p}{p-1}\right)^{p/2} \left(\frac{p-1}{2e}\right)^{1/2} =$$

$$\approx \sqrt{\frac{p-1}{2\pi}}$$

On the other hand, Using *Lemma* 1.4:

$$\int_\epsilon^1 \left(1-t^2\right)^{\frac{p-1}{2}} dt \leq \int_\epsilon^1 \exp -t^2 \frac{p-1}{2} dt$$

Using the following change of variables $t = s + \epsilon \Rightarrow dt = ds$. And the new integration limits are: $\epsilon = s + \epsilon \Rightarrow s = 0$ and $1 = s + \epsilon \Rightarrow s = 1 - \epsilon$, hence:

$$\int_\epsilon^1 \exp -t^2 \frac{p-1}{2} dt = \int_0^{1-\epsilon} \exp -(s+\epsilon)^2 \frac{p-1}{2} dt =$$

$$= \int_0^{1-\epsilon} \exp -\left(s^2 + \epsilon^2 + 2s\epsilon\right) \frac{p-1}{2} dt = \exp -\left(\epsilon^2 \frac{p-1}{2}\right) \int_0^{1-\epsilon} \exp -\left(s^2 + 2s\epsilon\right) \frac{p-1}{2} dt =$$

$$\leq \frac{1}{2\sqrt{p-1}} 2\pi \exp\left(-\epsilon^2 \frac{p-1}{2}\right)$$

Therefore, putting all the pieces together:

$$P\left\{(t_1, \ldots, t_p) \in B_p : |t_p| \geq \epsilon\right\} = 2\frac{\text{vol}[B_{p-1}]}{\text{vol}[B_p]} \int_\epsilon^1 \left(1-t^2\right)^{\frac{p-1}{2}} dt =$$

$$\leq 2\sqrt{\frac{p-1}{2\pi}} \frac{1}{2\sqrt{p-1}} 2\pi \exp\left(-\epsilon^2 \frac{p-1}{2}\right) = \sqrt{2\pi} \exp\left(-\epsilon^2 \frac{p-1}{2}\right)$$

$$\therefore P\left\{(t_1, \ldots, t_p) \in B_p : |t_p| \geq \epsilon\right\} \leq \sqrt{2\pi} \exp\left(-\epsilon^2 \frac{p-1}{2}\right)$$

And then:

$$P\left\{(t_1, \ldots, t_p) \in B_p : |t_p| \leq \epsilon\right\} = 1 - P\left\{(t_1, \ldots, t_p) \in B_p : |t_p| \geq \epsilon\right\}$$

$$\therefore P\left\{(t_1, \ldots, t_p) \in B_p : |t_p| \leq \epsilon\right\} \geq 1 - \sqrt{2\pi} \exp\left(-\epsilon^2 \frac{p-1}{2}\right)$$

Solution 1.4. Using the following script in Python:

(a)
```
import numpy as np
import matplotlib.pyplot as plt

for p in [10,100,1000]:
    x = np.random.normal(0,1, size = (100, p))
    l1 = np.linalg.norm(x, ord = 1, axis = 1)
    l2 = np.linalg.norm(x, axis = 1)
    plt.subplot(1,2,1)
    plt.hist(l1, bins = 'auto', ec = 'black', color = 'slateblue',density
                                            = True)
    plt.title('$l1$ norm ' + str(p))
    plt.xlabel('$x$')
    plt.ylabel('$\\rho(x)$')
    plt.subplot(1,2,2)
    plt.hist(l2, bins = 'auto', ec = 'black', color = 'slateblue',density
                                            = True)
    plt.title('$l2$ norm ' + str(p))
    plt.show()
```

- $p = 1$

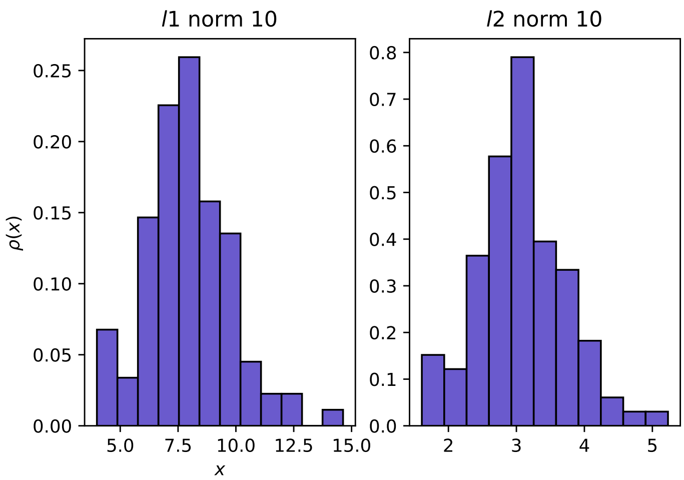

Figure 1: Distribution, $p = 1$, norm $l1$ and $l2$.

- $p = 100$

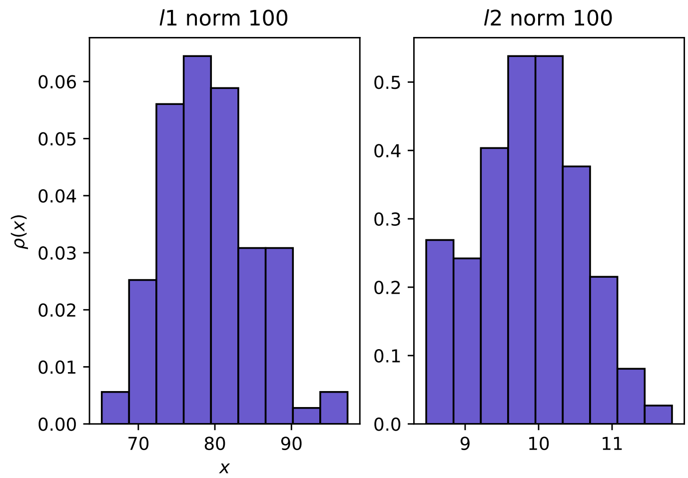

Figure 2: Distribution, $p = 100$, norm $l1$ and $l2$.

- $p = 1000$

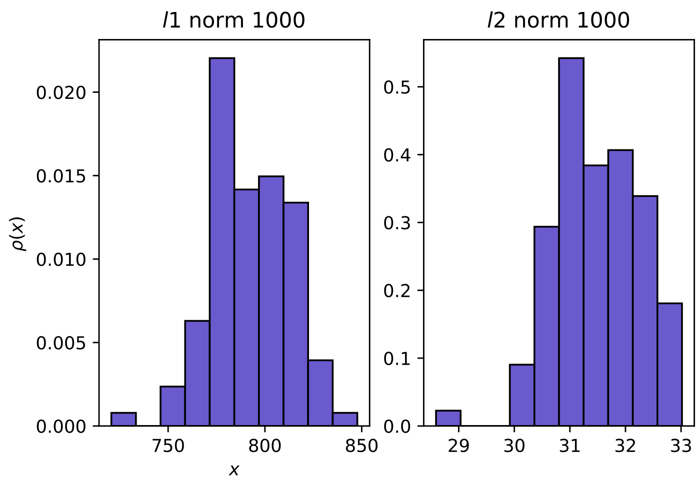

Figure 3: Distribution, $p = 1000$, norm $l1$ and $l2$.

(b)

```
def max_norm(x, **args):
    return np.max(np.abs(x), **args)

for p in [10,100,1000]:
    x = np.random.normal(0,1, size = (100, p))
    lmax = max_norm(x, axis =1)
    plt.hist(lmax, bins = 'auto', ec = 'black', color = 'slateblue',
                                          density = True)
    plt.title('$l$ max norm ' + str(p))
    plt.xlabel('$x$')
    plt.ylabel('$\\rho(x)$')
    plt.show()
```

- $p = 1$

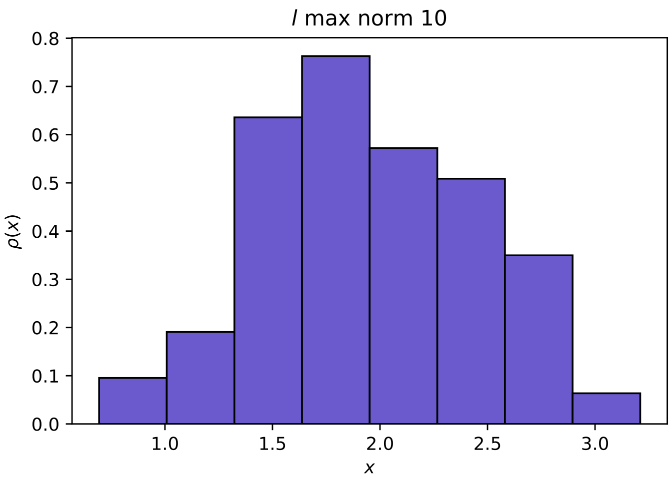

Figure 4: Distribution, $p = 1$, infinity norm.

- $p = 100$

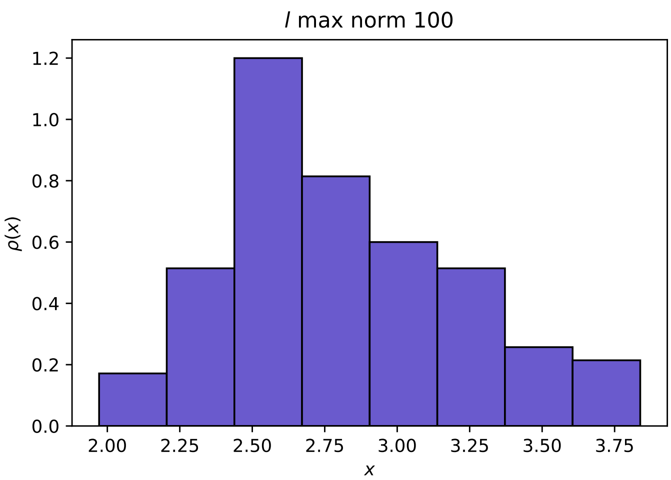

Figure 5: Distribution, $p = 100$, infinity norm.

- $p = 1000$

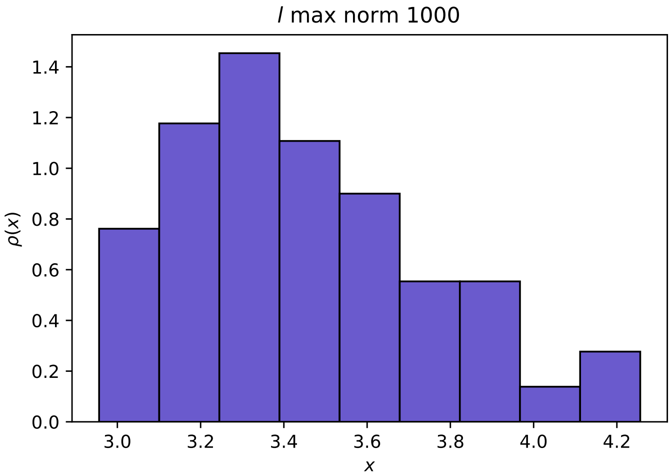

Figure 6: Distribution, $p = 1000$, infinity norm.

```
def estimate(ep, M, p):
    return np.sqrt(2/np.pi)*(p/((1+ep)*M)*np.exp(-0.5*((1+ep)**2)*M**2))

def mu(n,p):
    return (1/n)*(np.sqrt(n-1) + p)**2

def sigma(n,p):
    return mu(n,p)*np.float_power((1/np.sqrt(n-1))+(1/np.sqrt(p)),1/3)

for n in [1,10,50,100]:
dist = []
#m = mu(n = n,p = n)
#s = sigma(n,n)
for i in range(100):
    x = (1/n)*np.random.normal(0,1,size = (n,n))
    u, _ = np.linalg.eig(x@x.T)
    dist.append(max(u))
plt.hist(dist, bins = 'auto', ec = 'black', color = 'slateblue', density =
                                                                     True)
#plt.axvline(m, color = 'black', label = '$\mu$')
#plt.axvline(s, color = 'red', label = '$\sigma$')
plt.title('$f(X) = max(\lambda),$' + 'n = p =' +str(n))
plt.xlabel('$\lambda$')
plt.ylabel('$\\rho(\lambda)$')

 #plt.legend(loc = 'best')
plt.show()
```

(c)
- $p = n = 1$

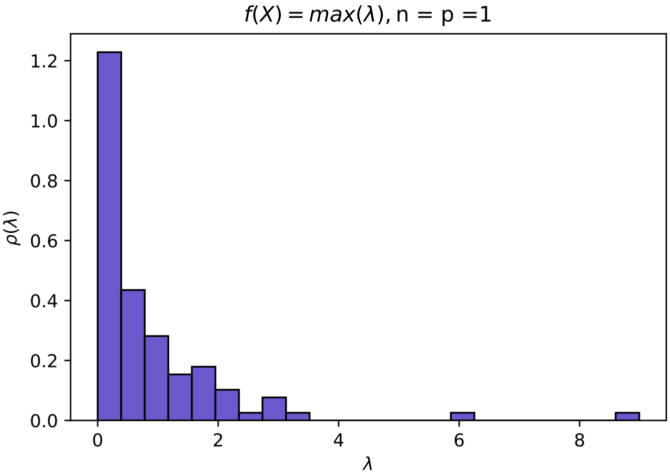

Figure 7: Distribution, $p = n = 1$.

- $p = n = 10$

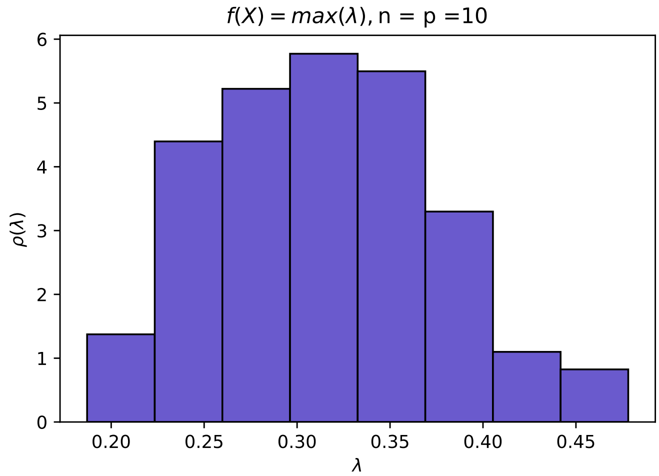

Figure 8: Distribution, $p = n = 10$.

- $p = n = 50$

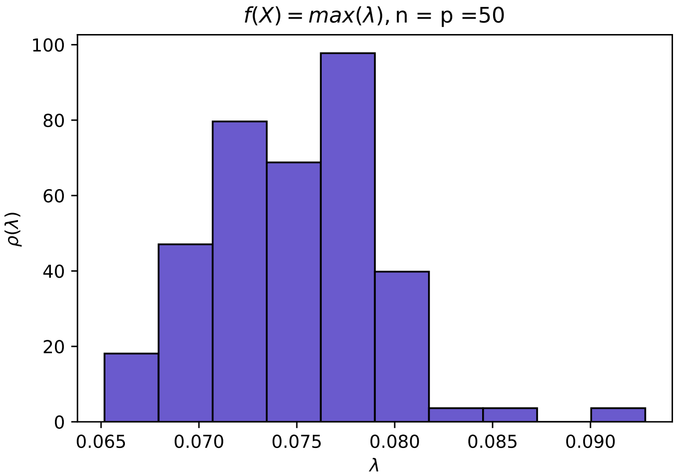

Figure 9: Distribution, $p = n = 50$.

- $p = n = 100$

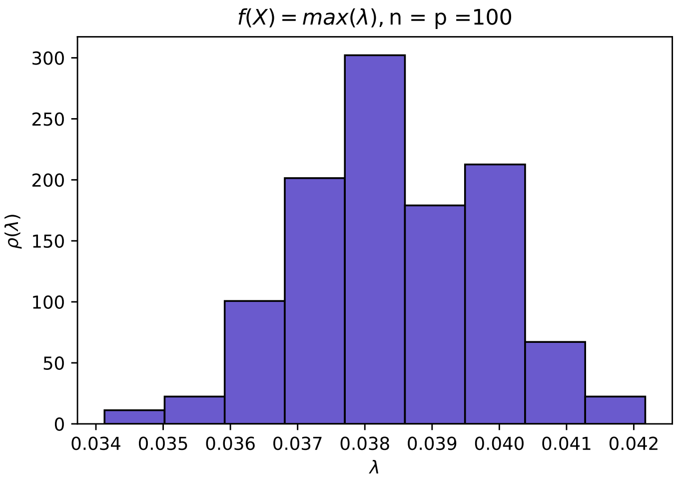

Figure 10: Distribution, $p = n = 100$.

2 Sheet 2 solutions

3 Sheet 3 solutions

4 Sheet 4 solutions

5 Sheet 5 solutions

6 Sheet 6 solutions