Generalization

สมาชิกกลุ่ม

นายกฤตพล ผ้าเจริญ 6401012620161

นายเจษฎา ศรีจุลโพธิ์ 6401012620170

นายศุภกร ผลศิริ 6401012620234

นายสิรภพ ห่วงวิไล 6401012630132

หาค่าความเอนเอียงและความแปรปรวนด้วย analytical method และ simulation ของแบบจำลองค่าคงที่และแบบจำลองเชิงเส้น แบบจำลองเชิงเส้นผ่านจุดกำเนิด

Sin(πx)

Analytical

Sin(πx) Constant model

$$g^{(D)}(x) = \frac{\sin(\pi x_1) + \sin(\pi x_2)}{2}$$

```
c_model = (sympy.sin(sympy.pi*x1) + sympy.sin(sympy.pi*x2)) / 2
```

```
g_bar: 0
bias^2: 0.500000000000000
variance: 0.250000000000000
E_over_dataset: 0.750000000000000
```

Sin(πx) Linear model

$$w_1^{(D)} = \frac{\sin(\pi x_1) - \sin(\pi x_2)}{x_1 - x_2} \qquad g^{(D)}(x) = (\sin(\pi x_2) - w_1^{(D)} x_2) + w_1^{(D)} x$$

```
slope = (sympy.sin(sympy.pi*x1) - sympy.sin(sympy.pi*x2)) / (x1 - x2)
model = (sympy.sin(sympy.pi*x2) - slope * x2) + slope * x # w0 + w1x
```

```
g_bar: 0.7759291740995 76*x
bias^2: 0.206716840217646
variance: 1.676282395045049
E_over_dataset: 1.882999235262 69
```

Sin(πx) Linear model through the origin

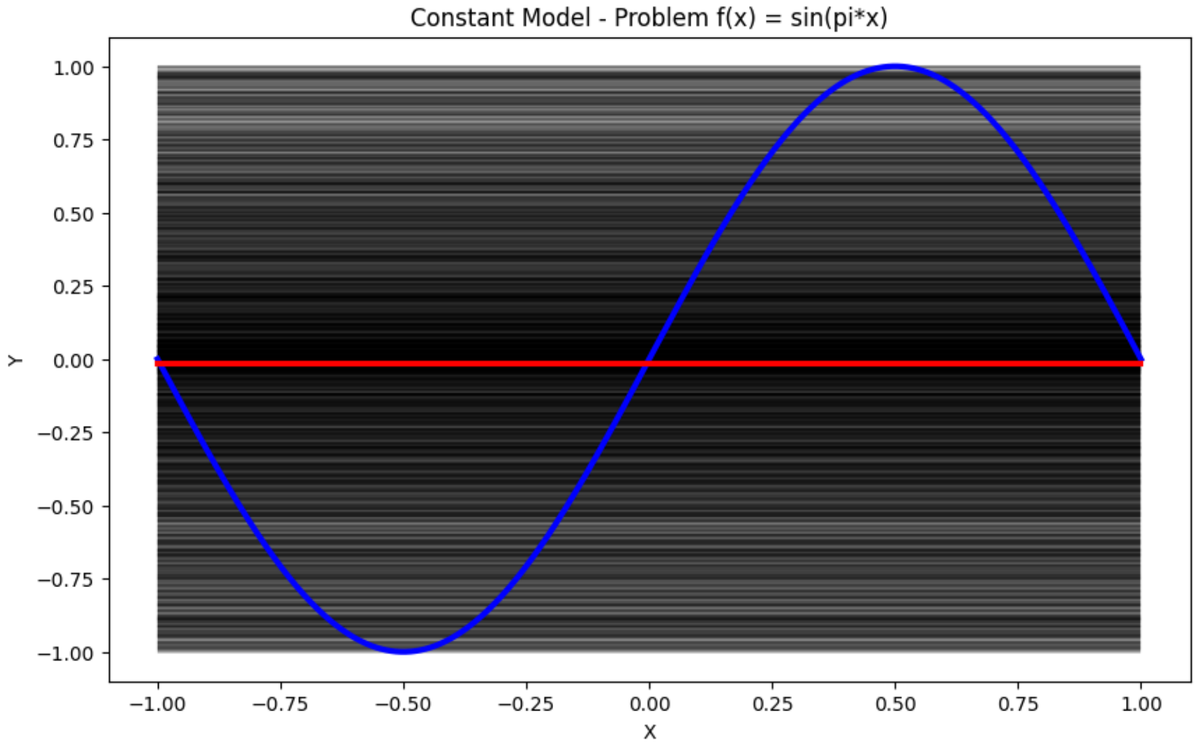
$$g^{(D)}(x) = \frac{(x_1 \sin(\pi x_1) + x_2 \sin(\pi x_2))x}{(x_1{}^2 + x_2{}^2)}$$

```
slope_sym = (((x1 * sympy.sin(sympy.pi*x1)) + (x2 * sympy.sin(sympy.pi*x2))) / (x1 ** 2 + x2 ** 2))
model = slope_sym * x
```

```
g_bar: 1.4280271748424 7*x
bias^2: 0.270643535713256
variance: 0.236575925913684
E_over_dataset: 0.507219461626941
```

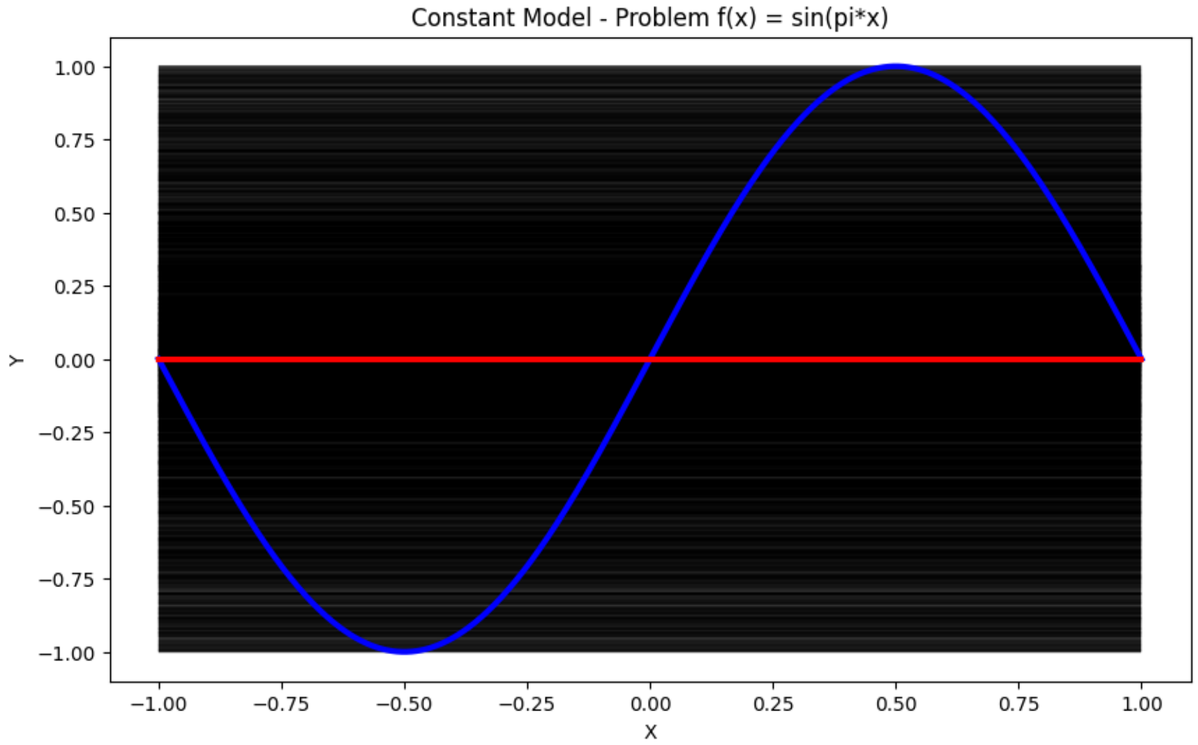
Simulation

Sin(πx) Constant model



samples = 5000

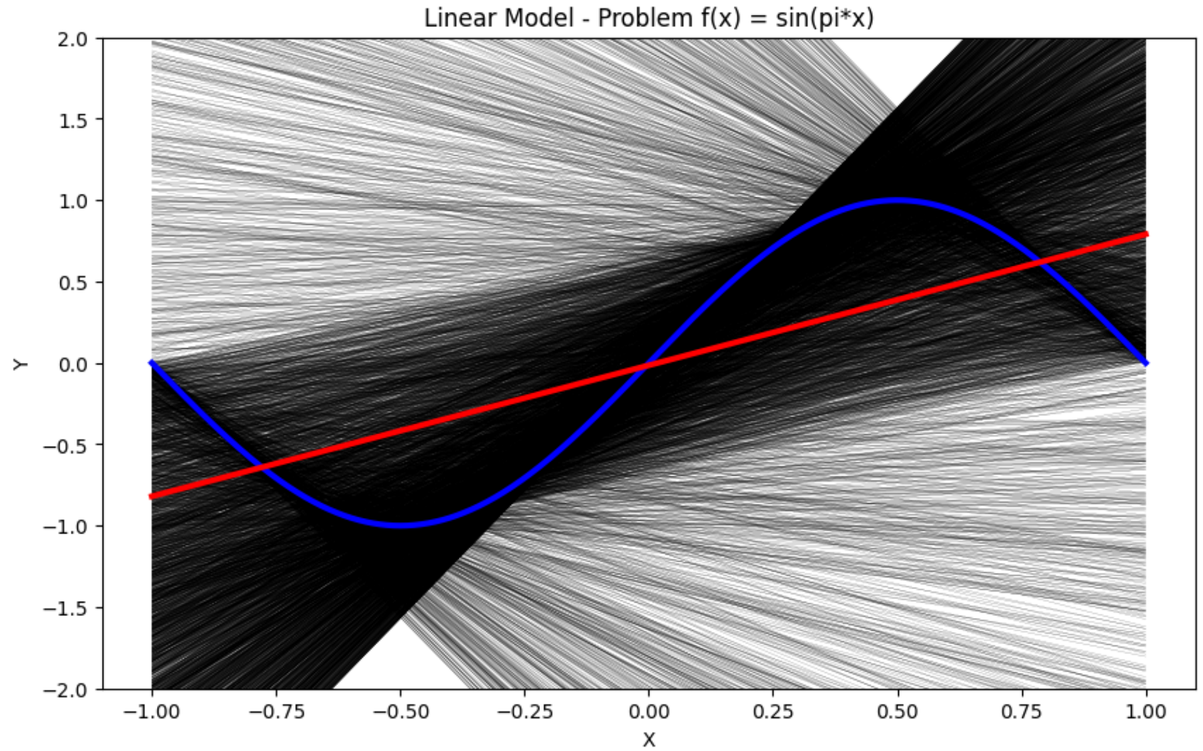
```
var: 0.2502259754968540 5
bias: 0.4997549550103806 4
```



samples = 10000

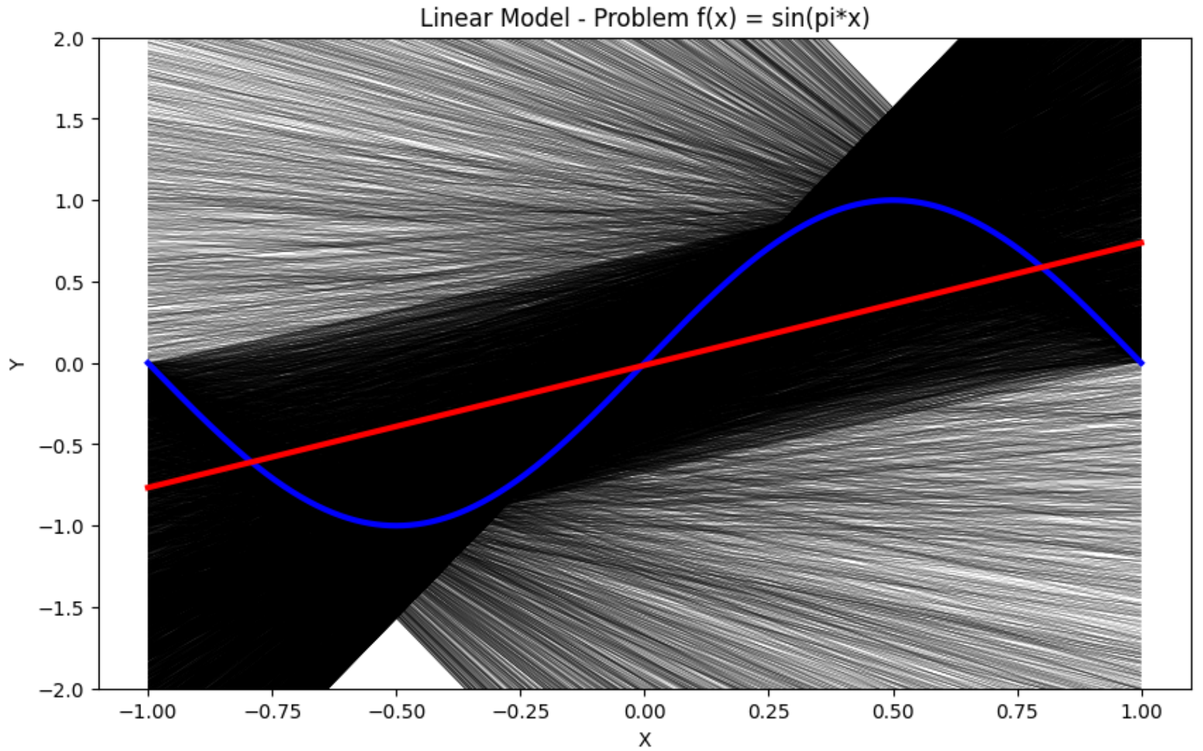
```
var: 0.254145724684096
bias: 0.4995089583658492 4
```

Sin(πx) Linear model



samples = 5000

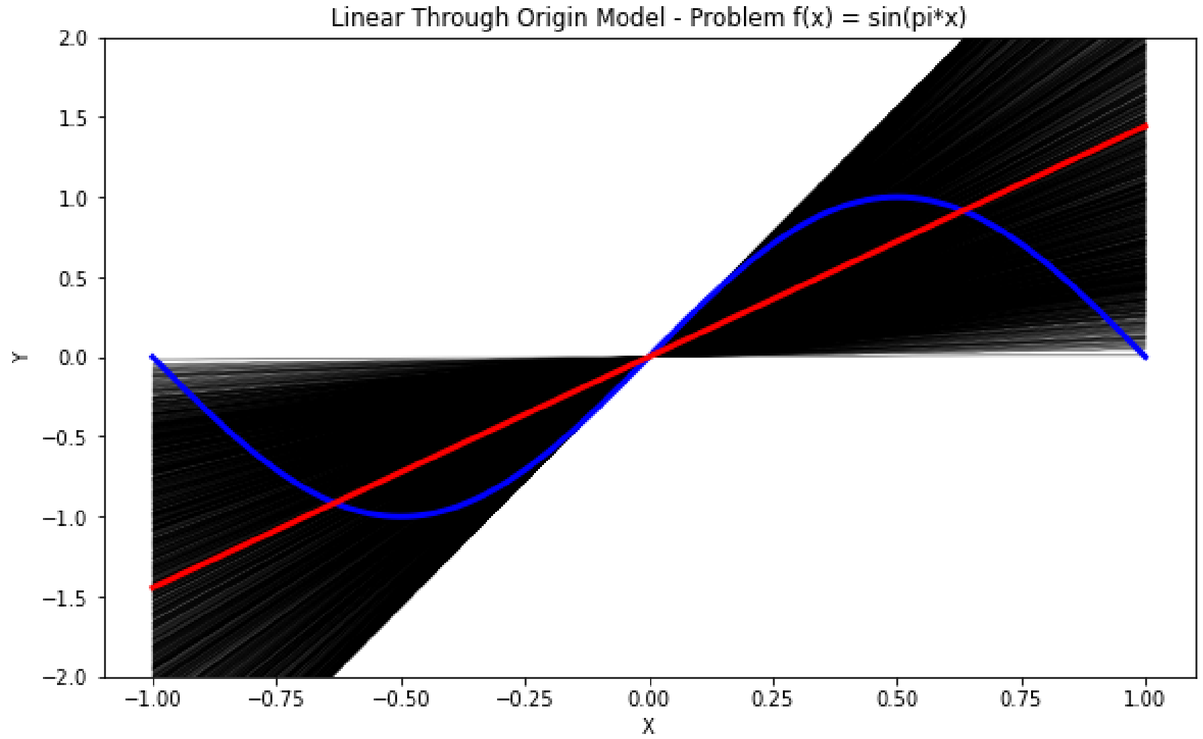
```
var: 1.6534596752752684
bias: 0.20809490494658292
```



samples = 10000

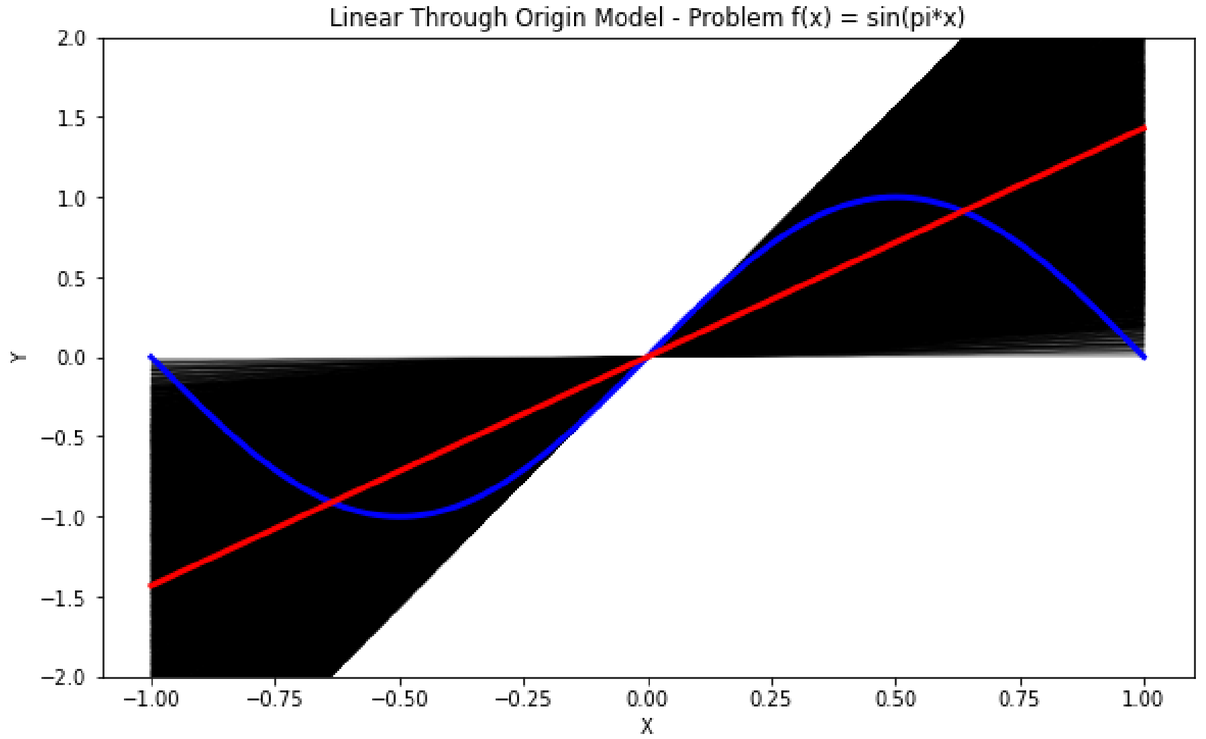
```
var: 1.7000135843316202
bias: 0.210334212362329
```

Sin(πx) Linear model through the origin



samples = 5000

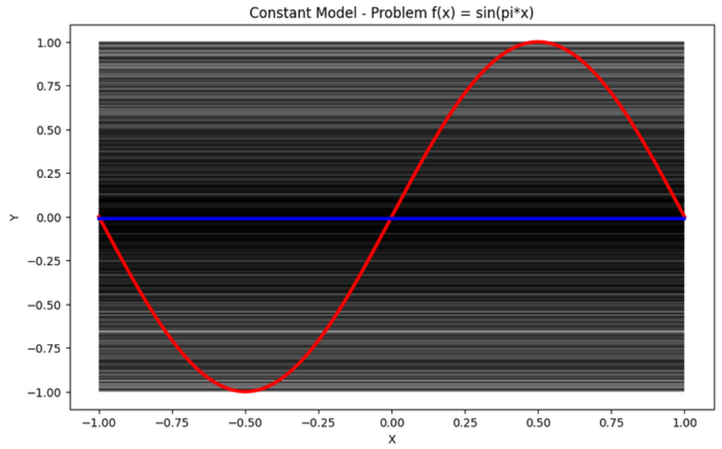
```
var: 0.23660925489278953
bias: 0.27756502957374346
```

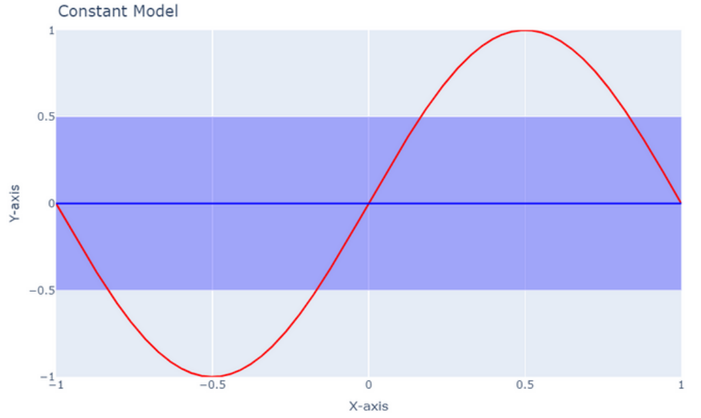


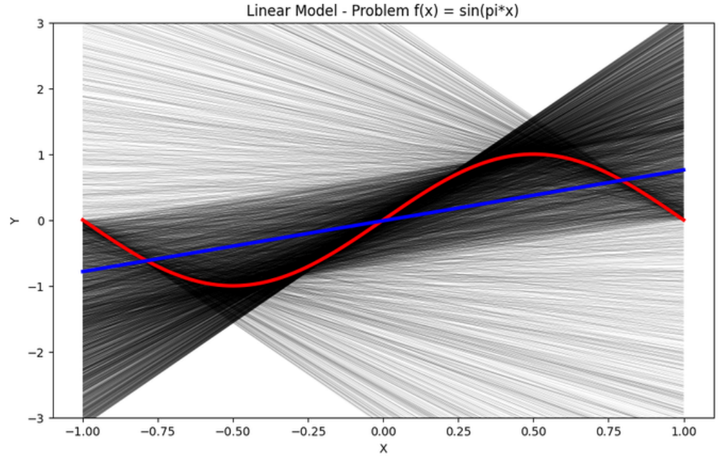
samples = 10000

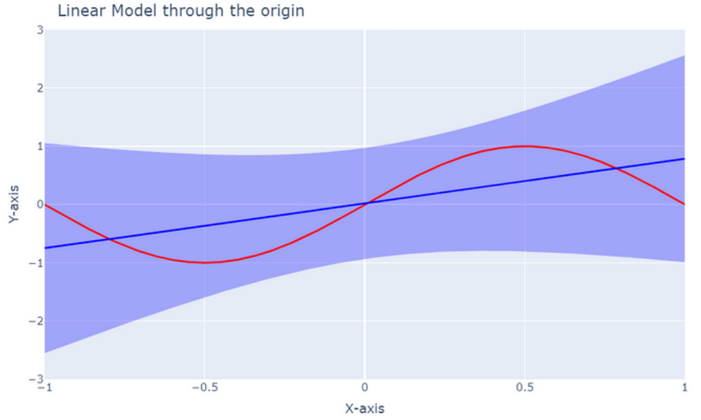
```
var: 0.2351656468190997
bias: 0.27352453650918174
```

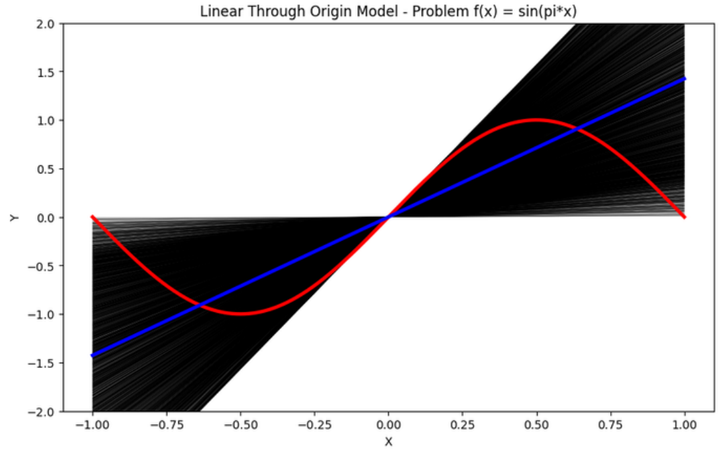
Sin(πx)	Variance	Bias
Constant	~0.25	~0.5
Linear	~1.7	~0.21
Linear through the origin	~0.235	~0.27

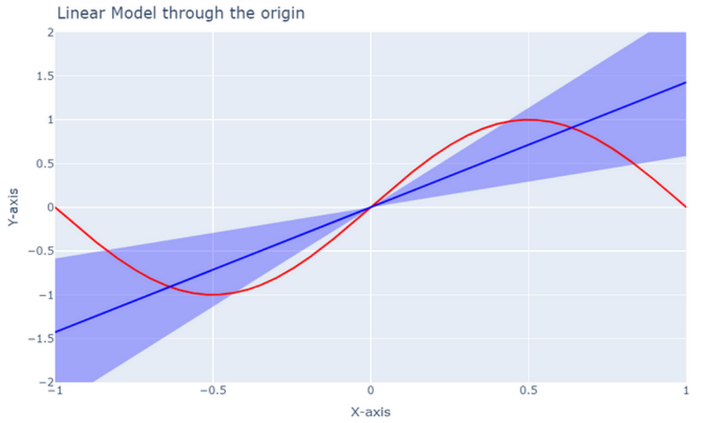












$$x^2$$

Analytical

x² Constant model

$$g^{(D)}(x) = \frac{x_1{}^2 + x_2{}^2}{2}$$

```
c_model = (x1**2 + x2**2) / 2
```

```
g_bar: 0.333333333333333
bias^2: 0.088888888888889
variance: 0.044444444444445
E_over_dataset: 0.133333333333333
```

x² Linear model

$$w_1^{(D)} = \frac{x_1^2 - x_2^2}{x_1 - x_2} \qquad g^{(D)}(x) = (x_2^2 - w_1^{(D)} x_2) + w_1^{(D)} x$$

```
slope = (x1**2 - x2**2) / (x1 - x2)
model = (x2**2 - slope * x2) + slope * x
```

```
g_bar: 0
bias^2: 0.200000000000000
variance: 0.333333333333333
E_over_dataset: 0.533333333333333
```

x² Linear model through the origin

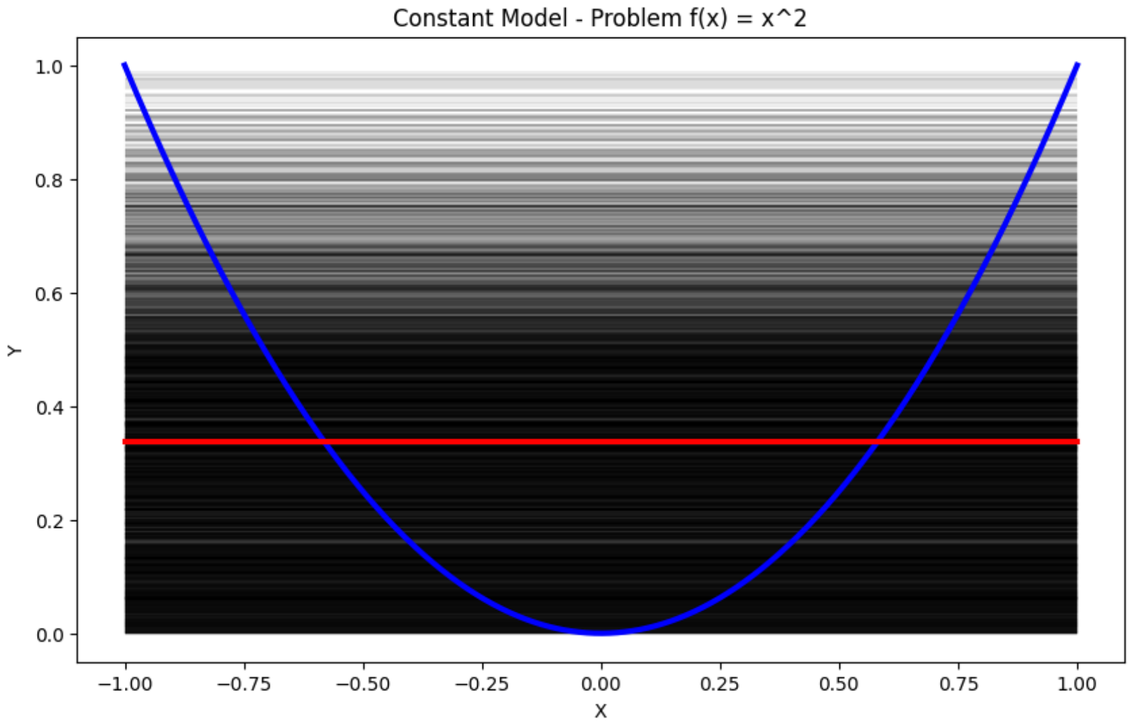
$$g^{(D)}(x) = \frac{(x_1{}^3 + x_2{}^3)x}{(x_1{}^2 + x_2{}^2)}$$

```
model = (((x1 ** 3) + (x2 **3)) * x / (x1 ** 2 + x2 ** 2))
```

```
g_bar: 0
bias^2: 0.200000000000000
variance: 0.114921303921007
E_over_dataset: 0.314921303921007
```

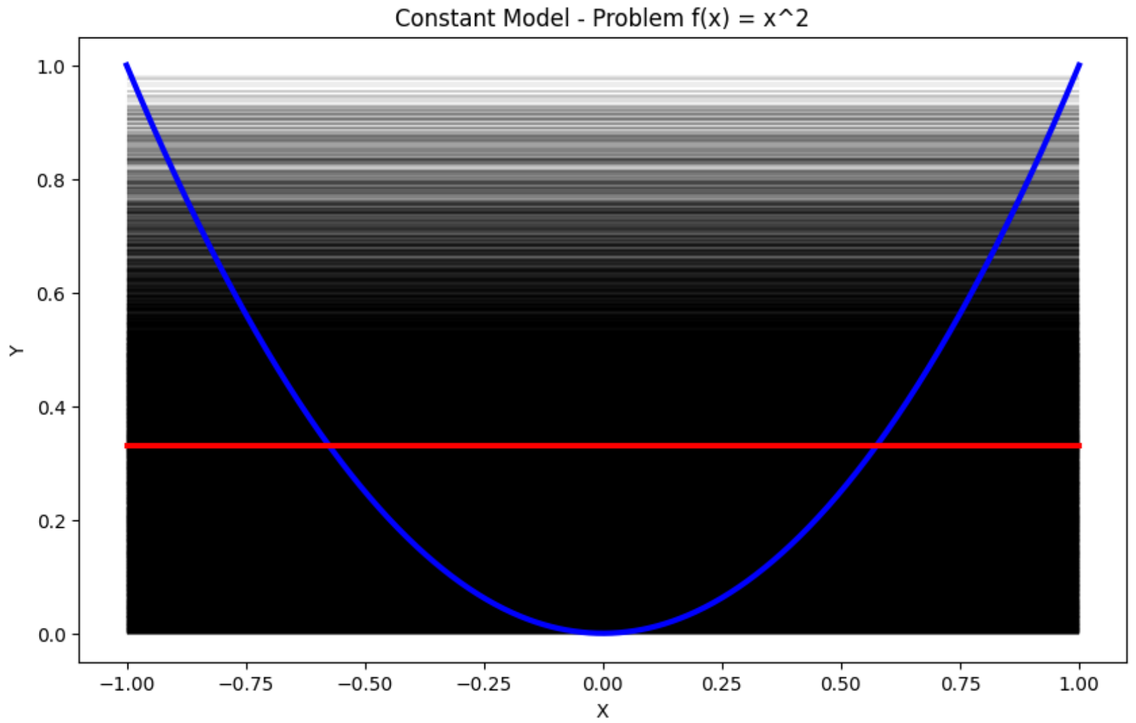
Simulation

x^2 Constant model



n = 5000

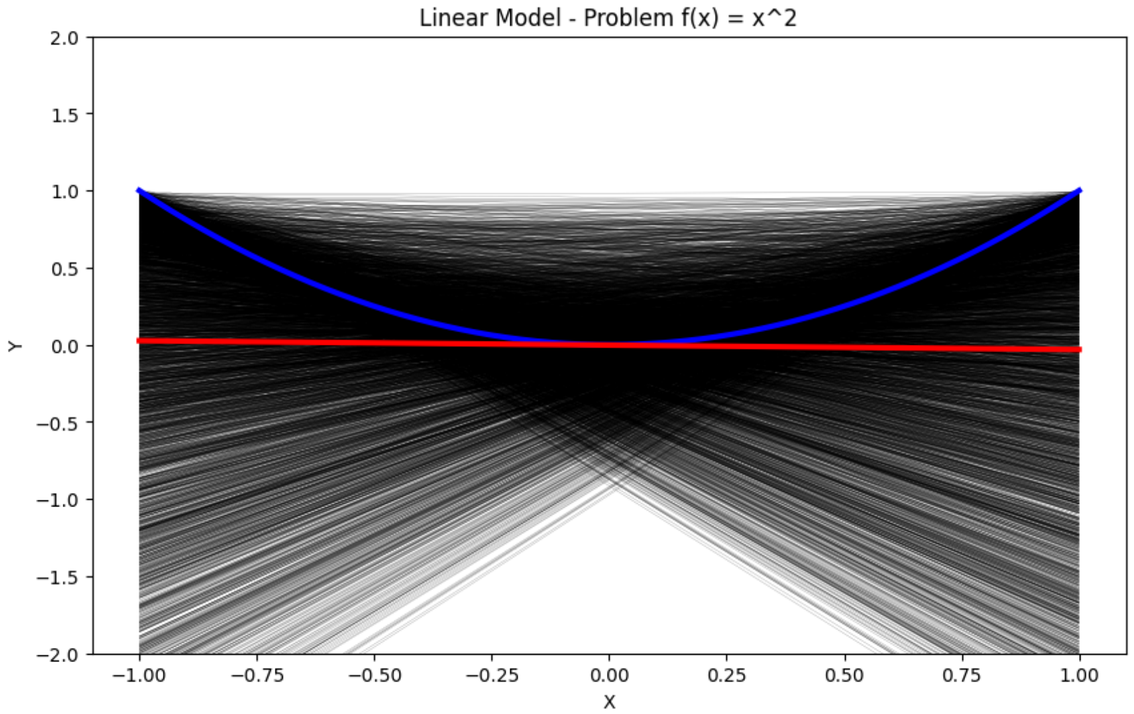
```
var: 0.04601692406217751 4
bias: 0.08925223262184732
```



n = 10000

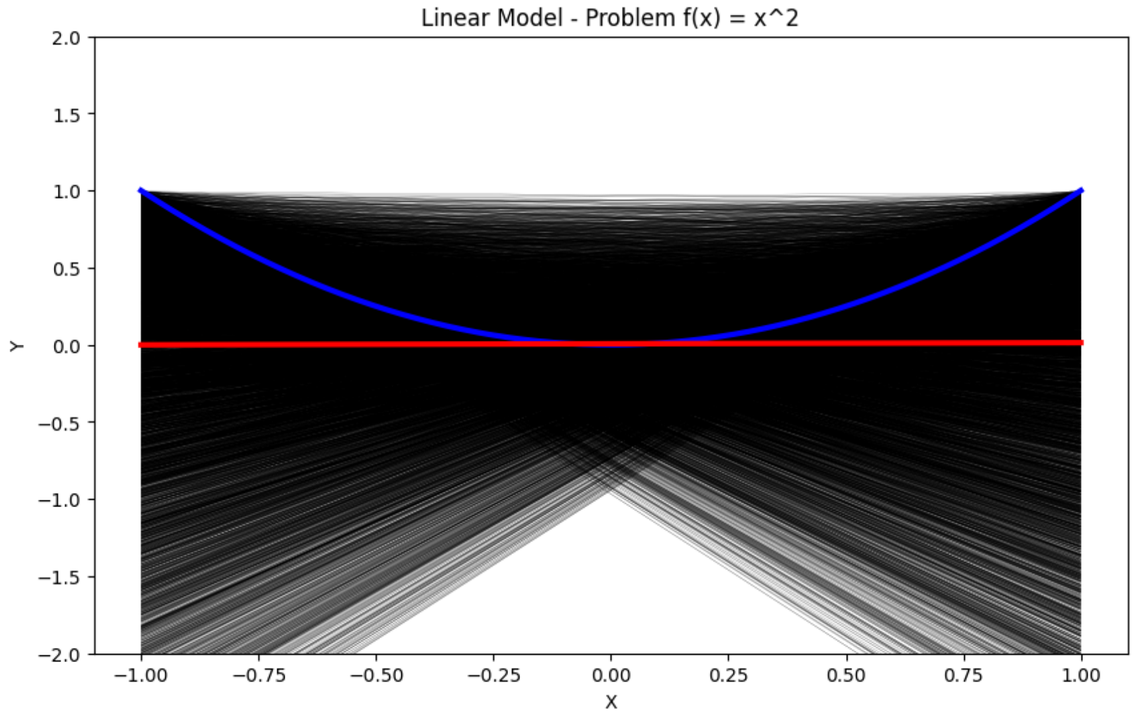
```
var: 0.04445679843037777
bias: 0.08925883967885433
```

x^2 Linear model



n = 5000

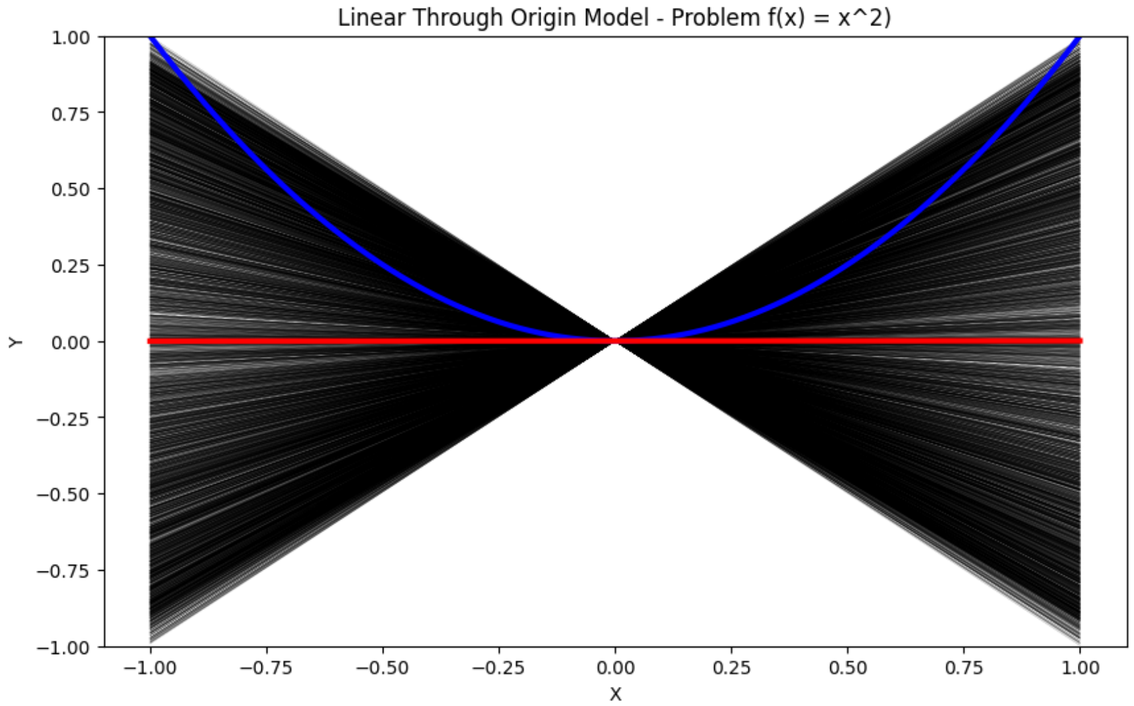
```
var: 0.33357440775814656
bias: 0.19734567965440977
```



n = 10000

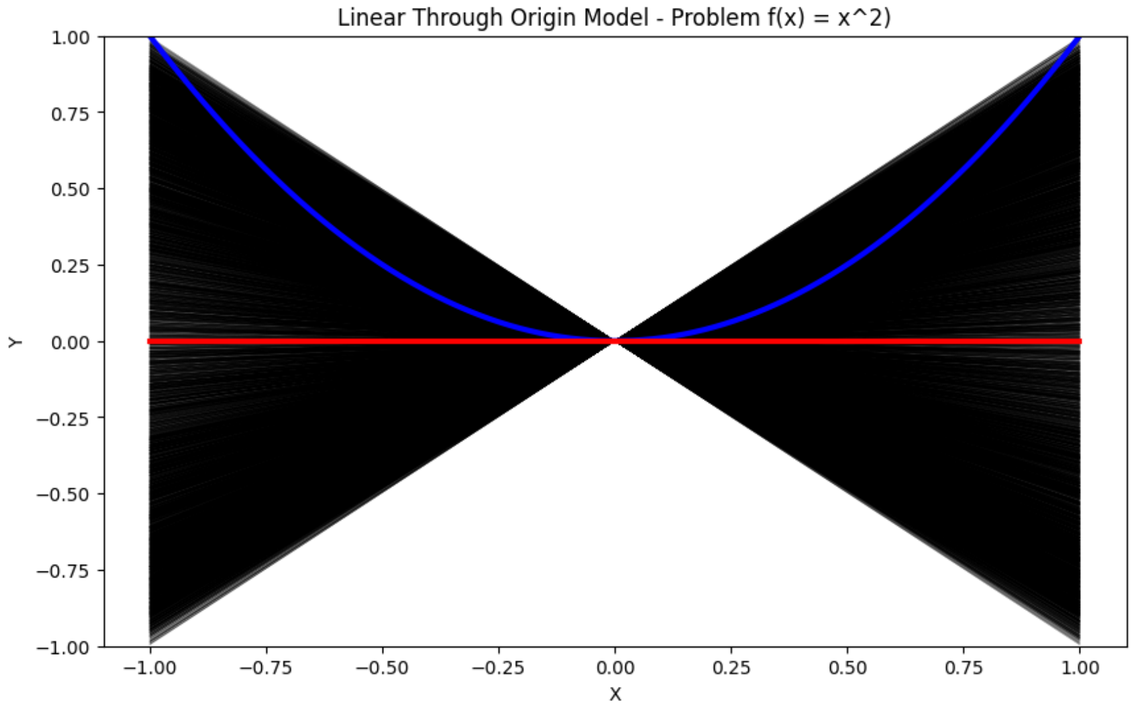
```
var: 0.3301366731765763
bias: 0.1963532523275827 2
```

x^2 Linear model through the origin



n = 5000

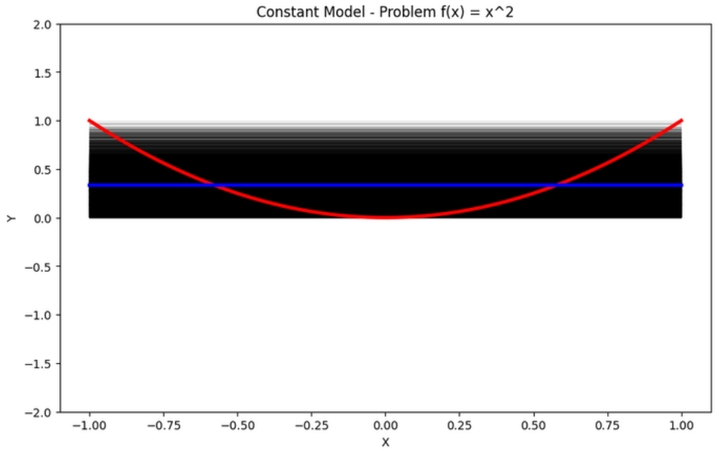
```
var: 0.11531590587989149
bias: 0.20084610334827863
```

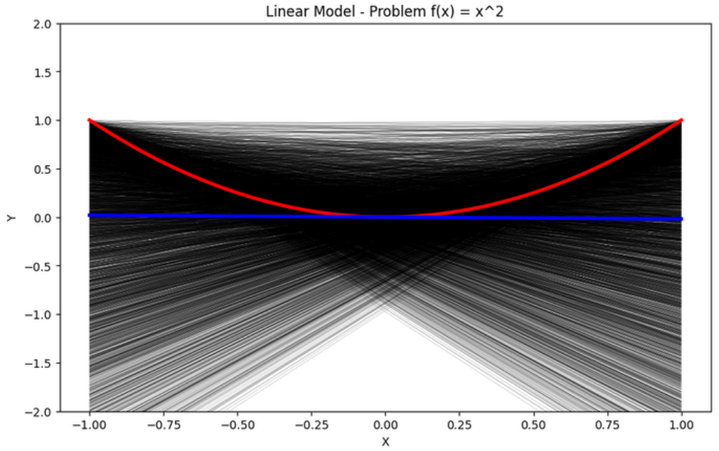


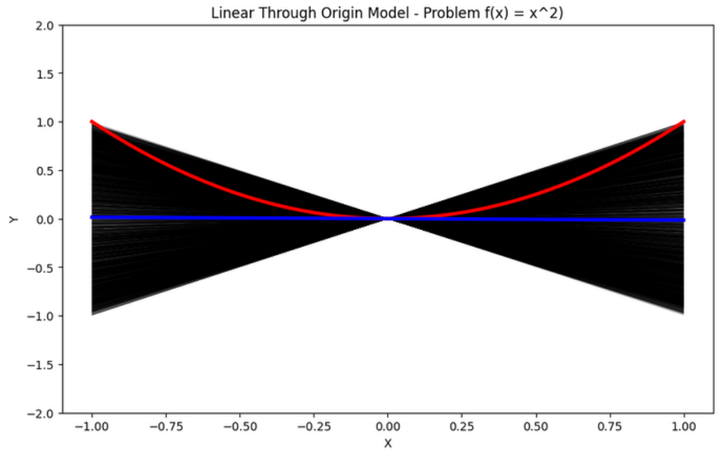
n = 10000

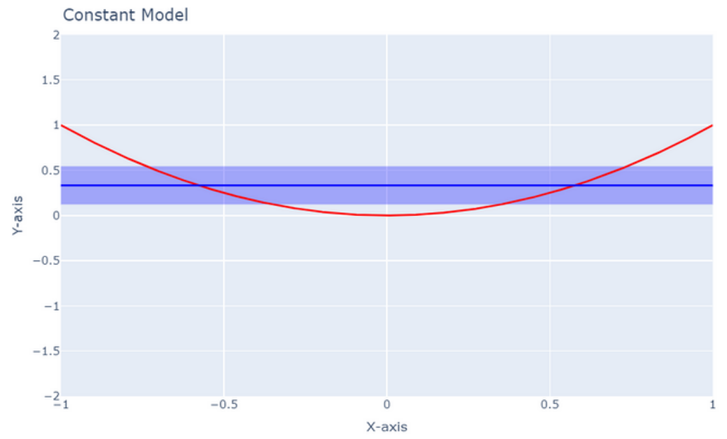
```
var: 0.11520616176500006
bias: 0.20080133512102438
```

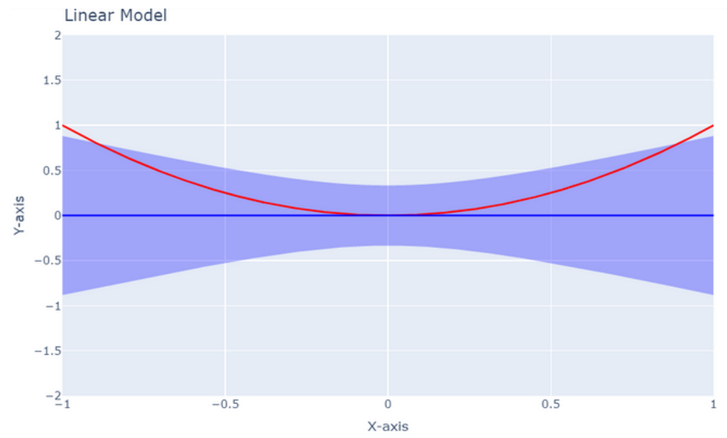
x^2	Variance	Bias
Constant	~0.045	~0.089
Linear	~0.33	~0.2
Linear through the origin	~0.115	~0.2

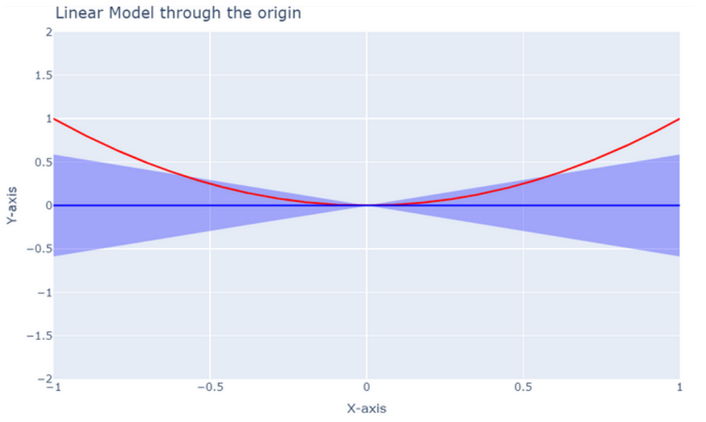






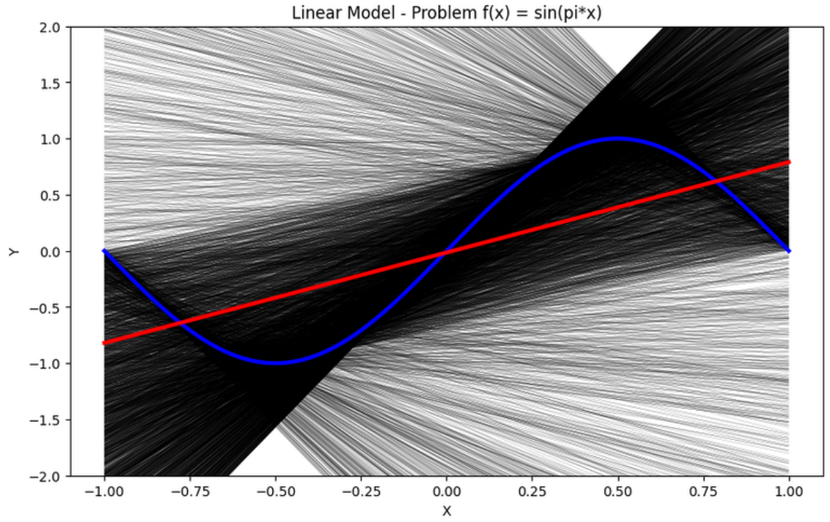






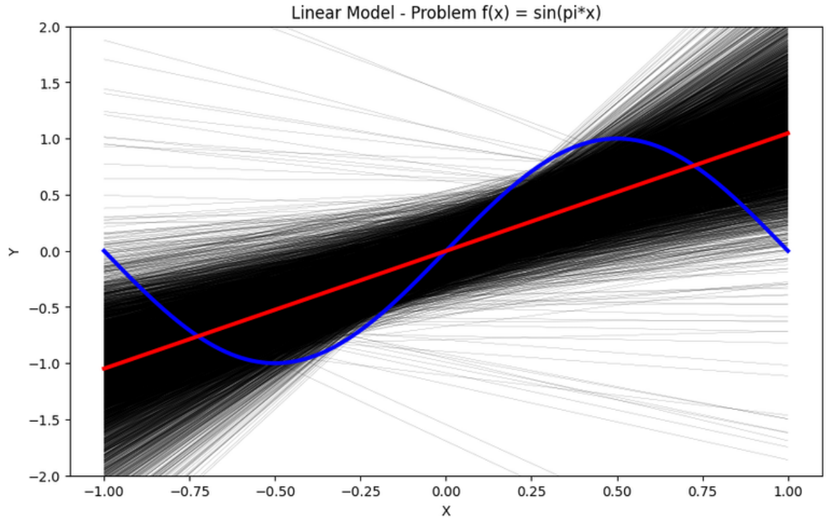
เขียนโปรแกรมสำหรับเส้นโค้งการเรียนรู้เปรียบเทียบระหว่างแบบจำลองค่าคงที่ แบบจำลองเชิงเส้น แบบจำลองเชิงเส้นผ่านจุดกำเนิด
และทดลองเพิ่มเติมด้วยการใส่สัญญาณรบกวน

Noiseless



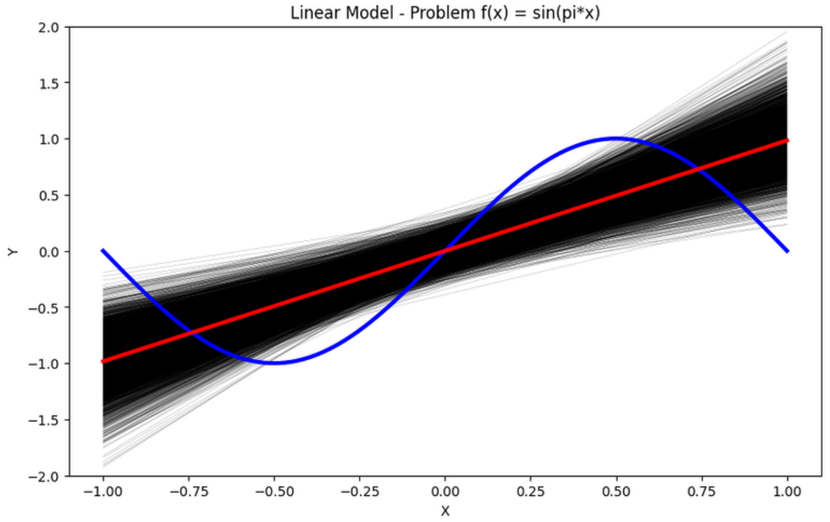
n = 2
samples = 5000

```
var: 1.6534596752752684
bias: 0.2080949049465829 2
```



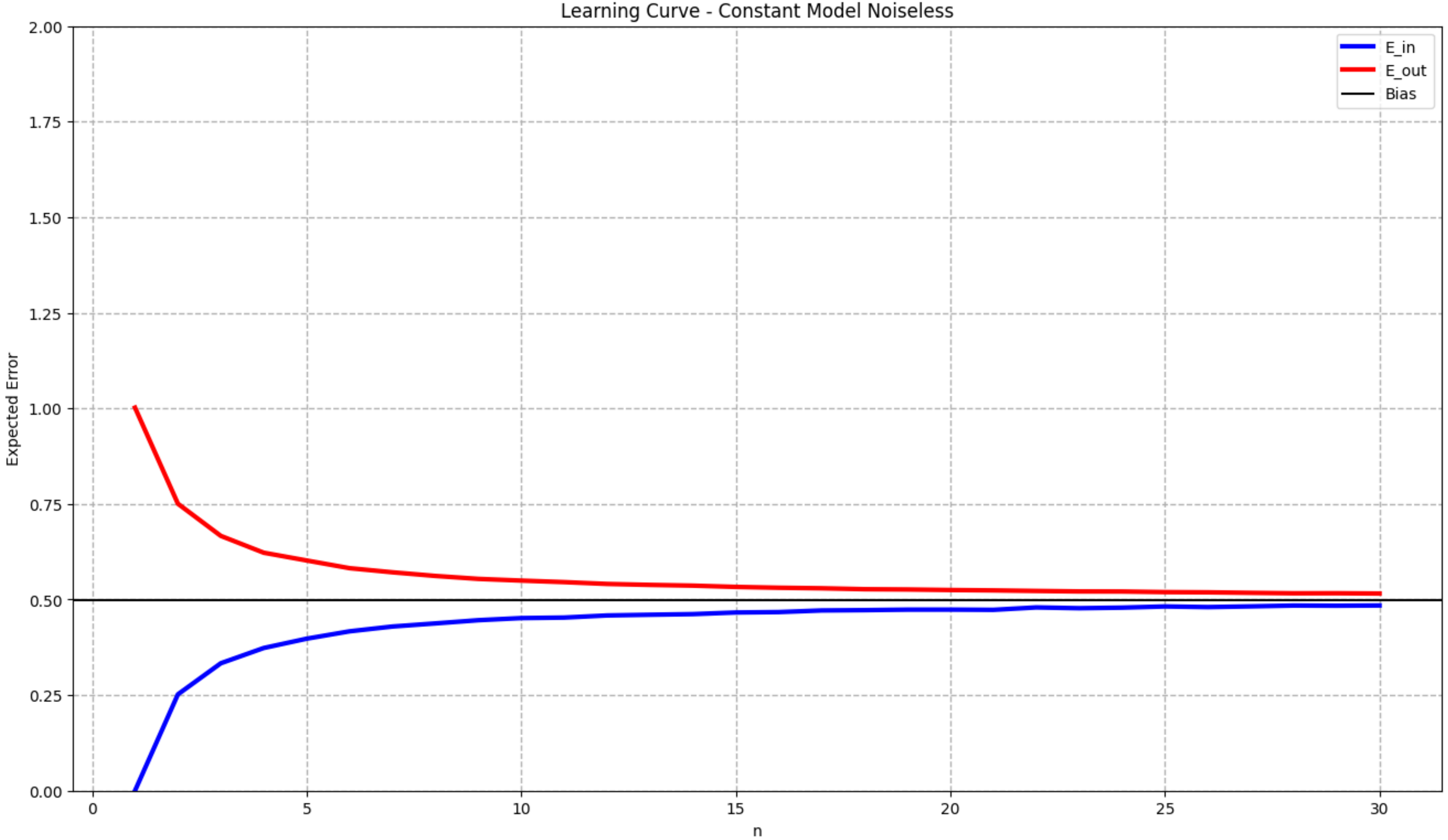
n = 7
samples = 5000

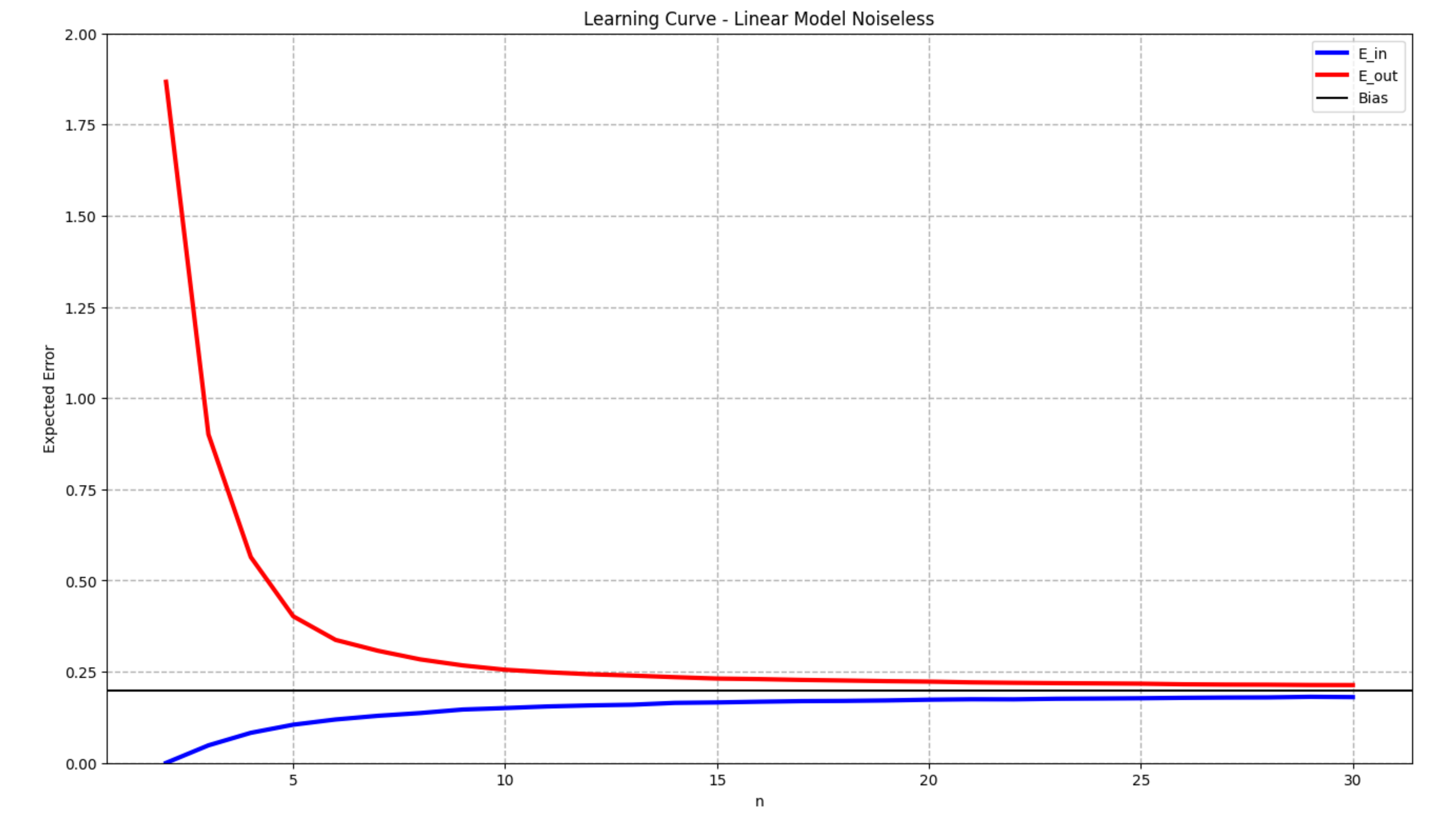
```
var: 0.09711397046160931
bias: 0.19980410355426487
```

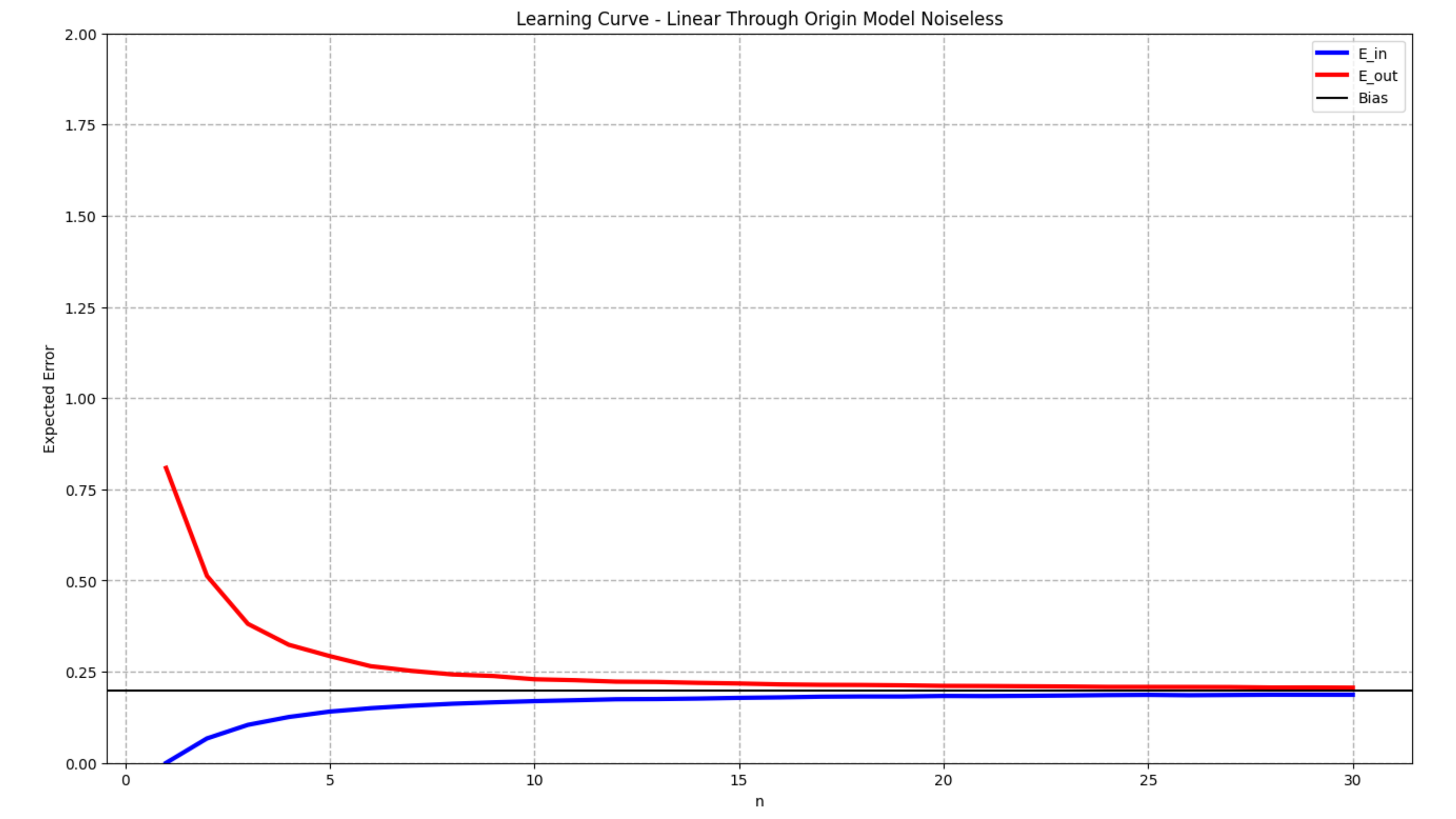


n = 20
samples = 5000

```
var: 0.025205299257163655
bias: 0.19707585124497692
```

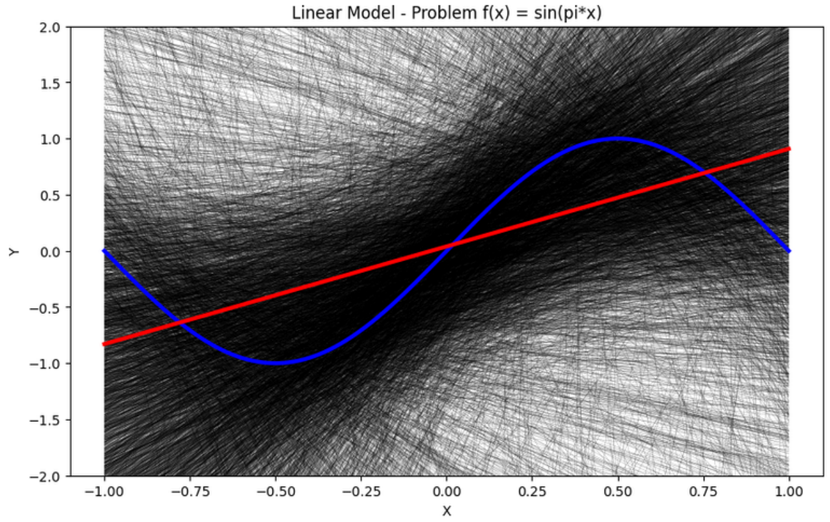

Learning Curve - Constant Model Noiseless
E_in
E_out
Bias
Expected Error
n
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
0
5
10
15
20
25
30

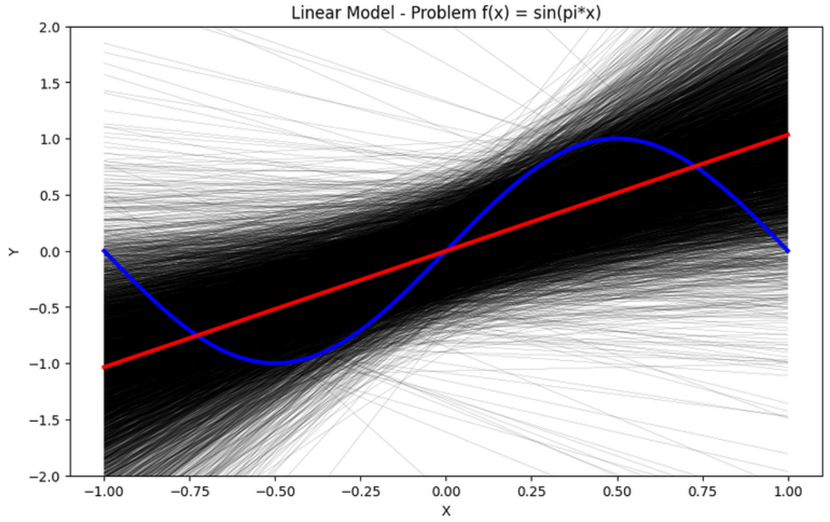

Learning Curve - Linear Model Noiseless
E_in
E_out
Bias
Expected Error
n
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
5
10
15
20
25
30

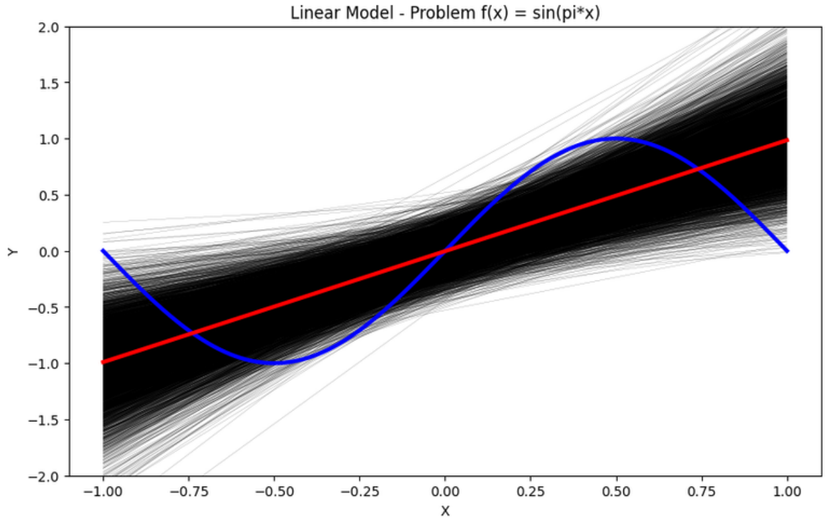

Learning Curve - Linear Through Origin Model Noiseless
E_in
E_out
Bias
Expected Error
n
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
0
5
10
15
20
25
30

ทดลองเพิ่มเติมด้วยการใส่สัญญาณรบกวน
ที่เป็น Normal Distribution Random
มี Standard deviation = 0.5
และ Mean = 0

Noisy







n = 2
samples = 5000

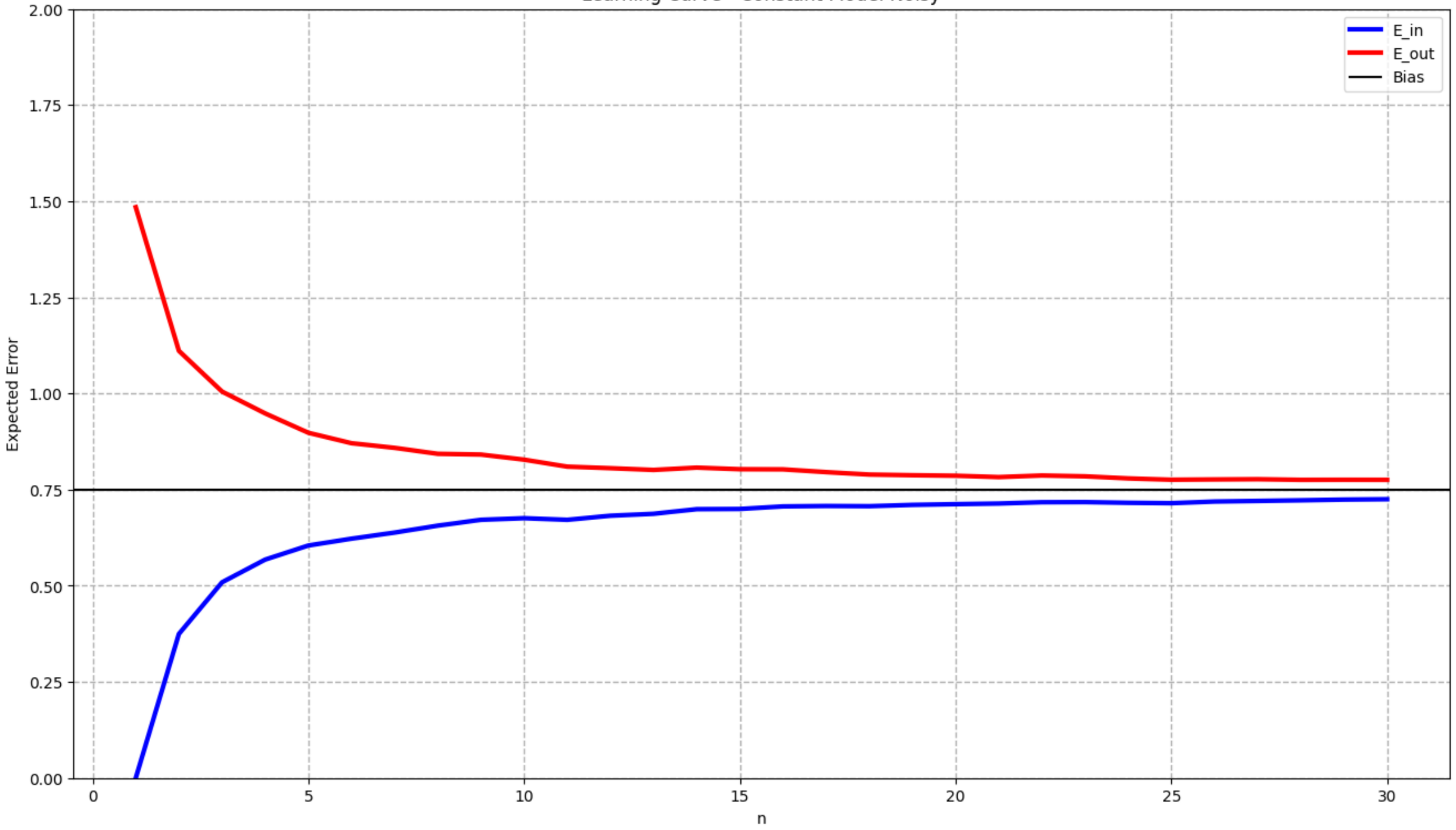
n = 7
samples = 5000

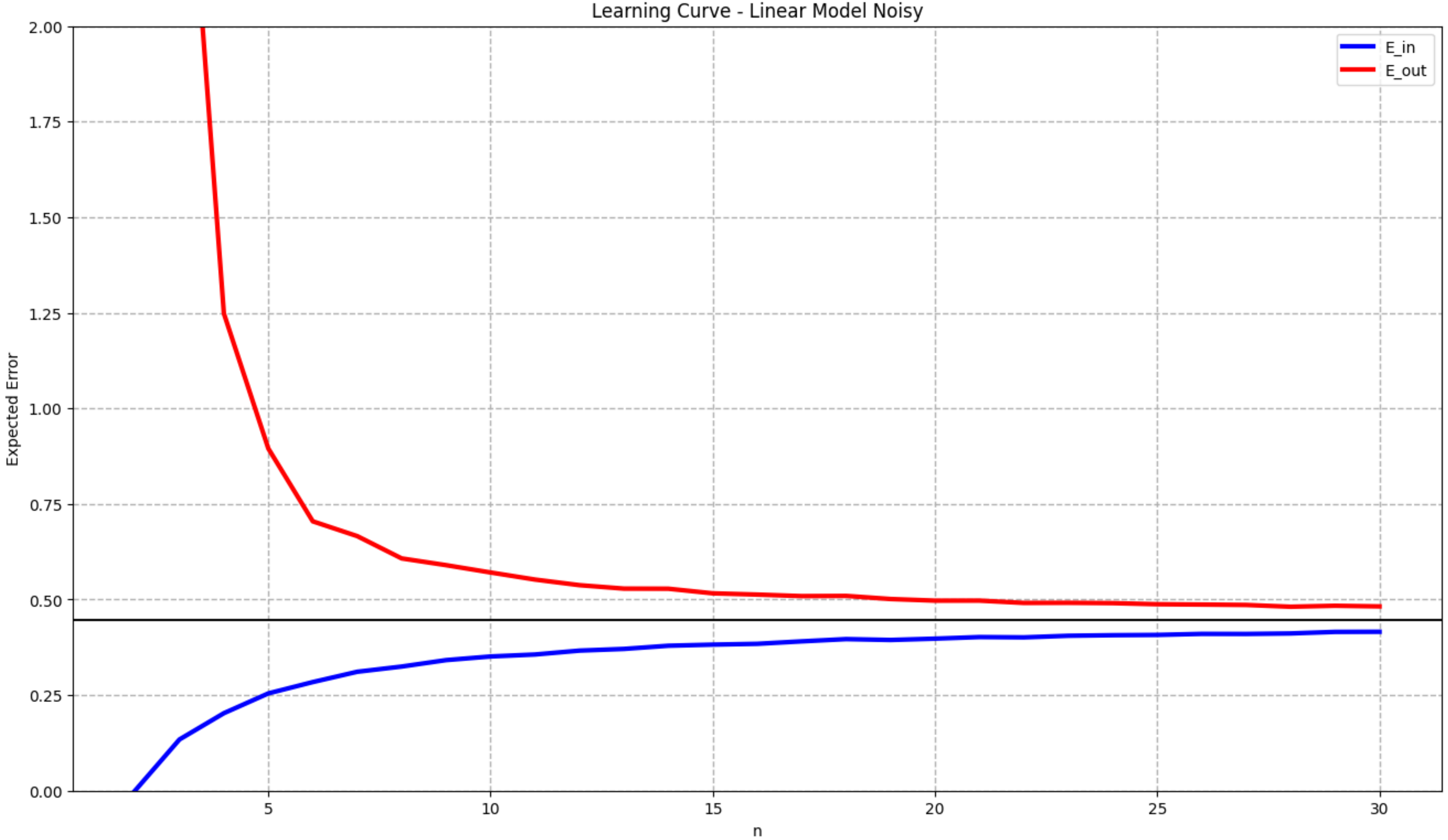
n = 20
samples = 5000

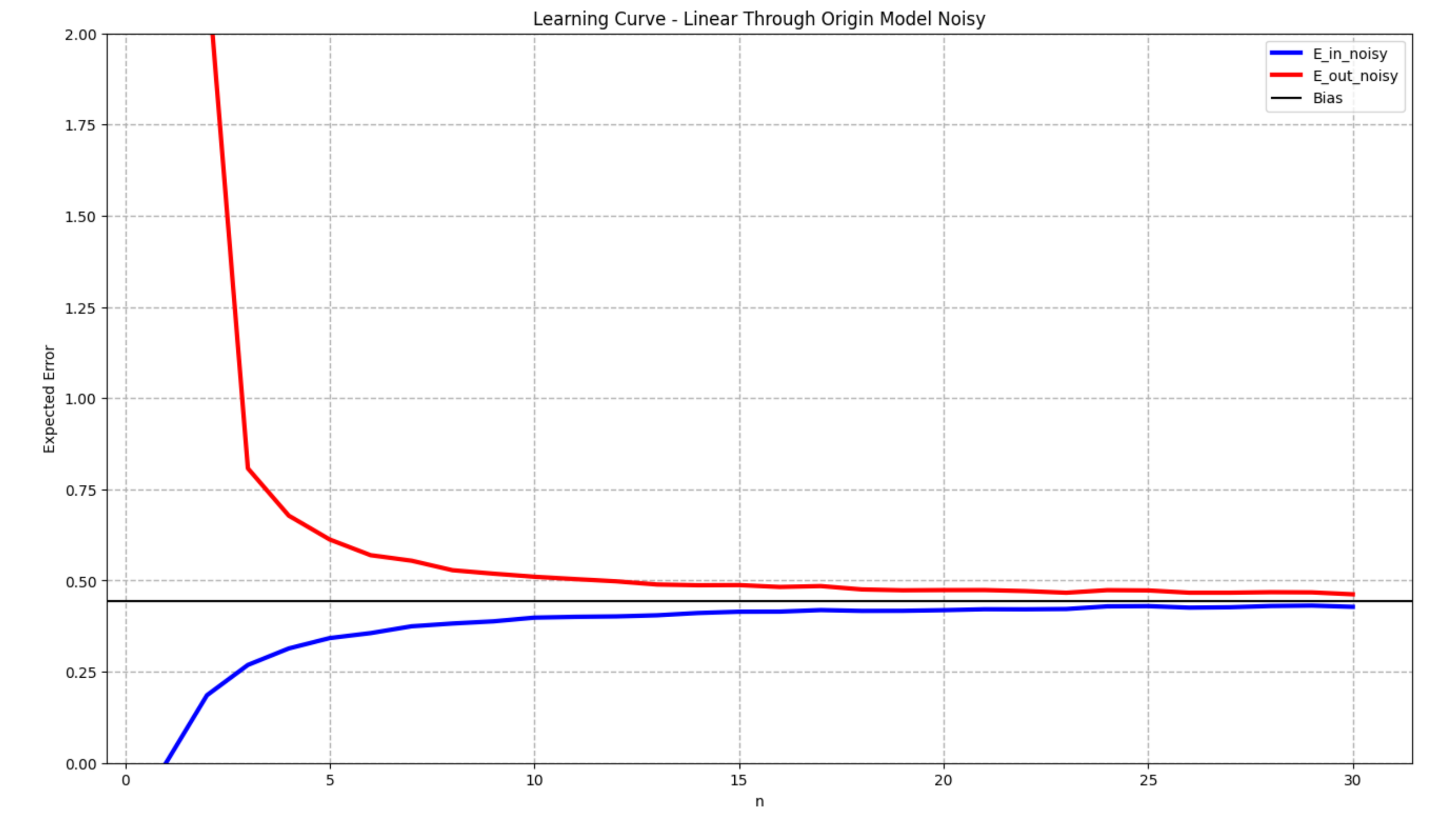
```
var: 490.5888579094394
bias: 0.44117389666702755
```

```
var: 0.2079979969217308
bias: 0.4576049647359865
```

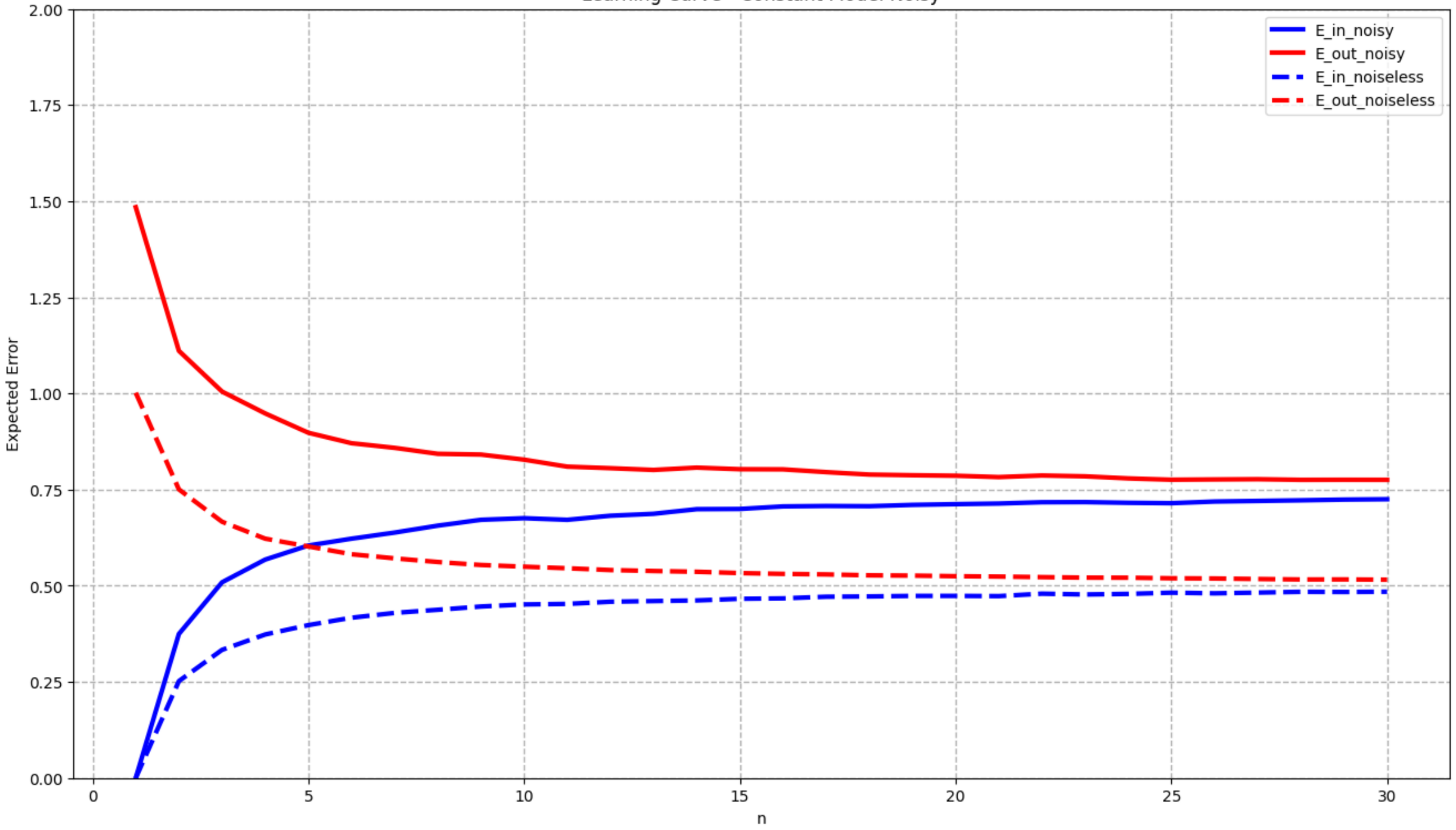
```
var: 0.05233605666666683055
bias: 0.4541831466756281
```

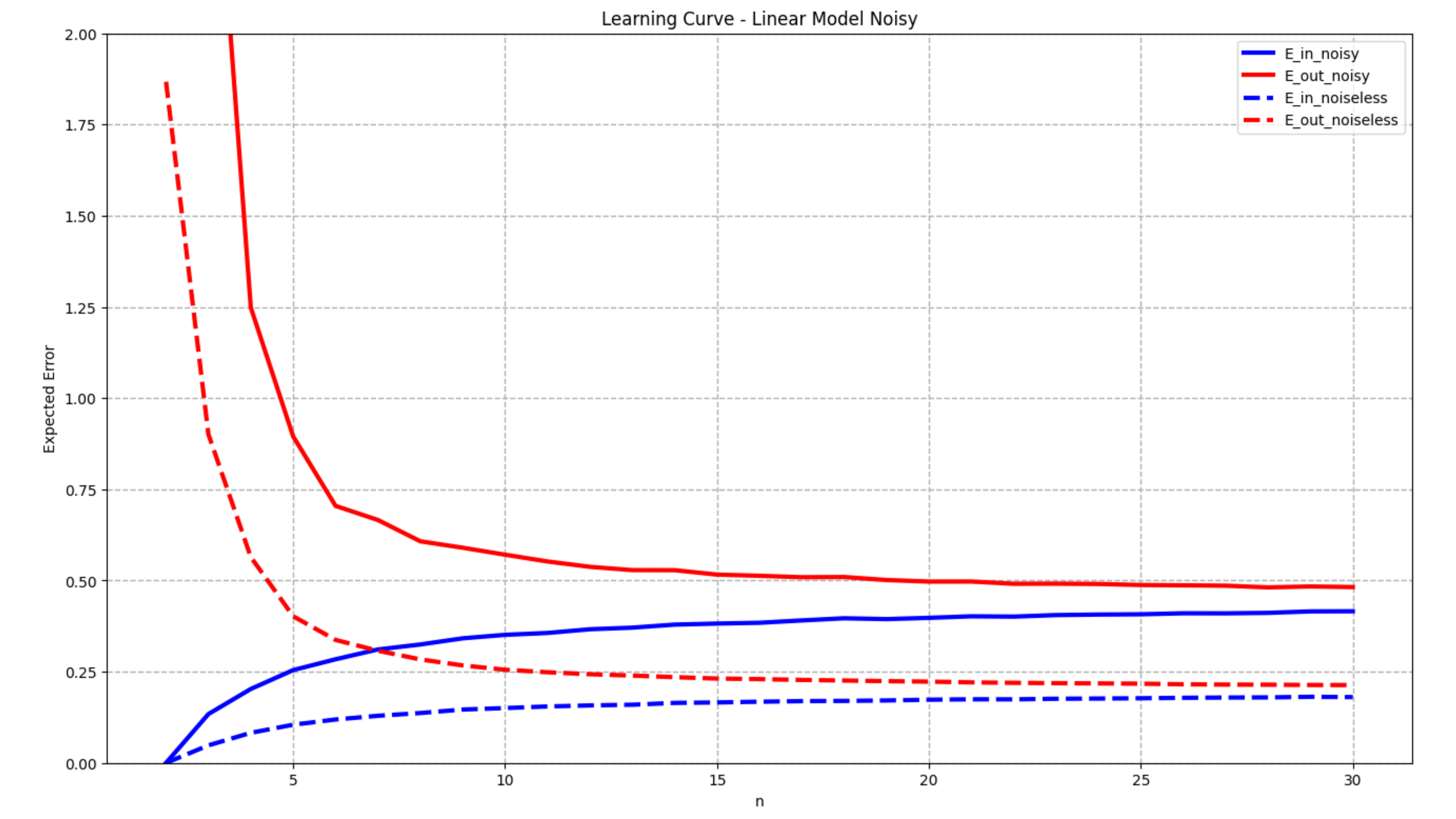

Learning Curve - Constant Model Noisy
E_in
E_out
Bias
Expected Error
n
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
0
5
10
15
20
25
30

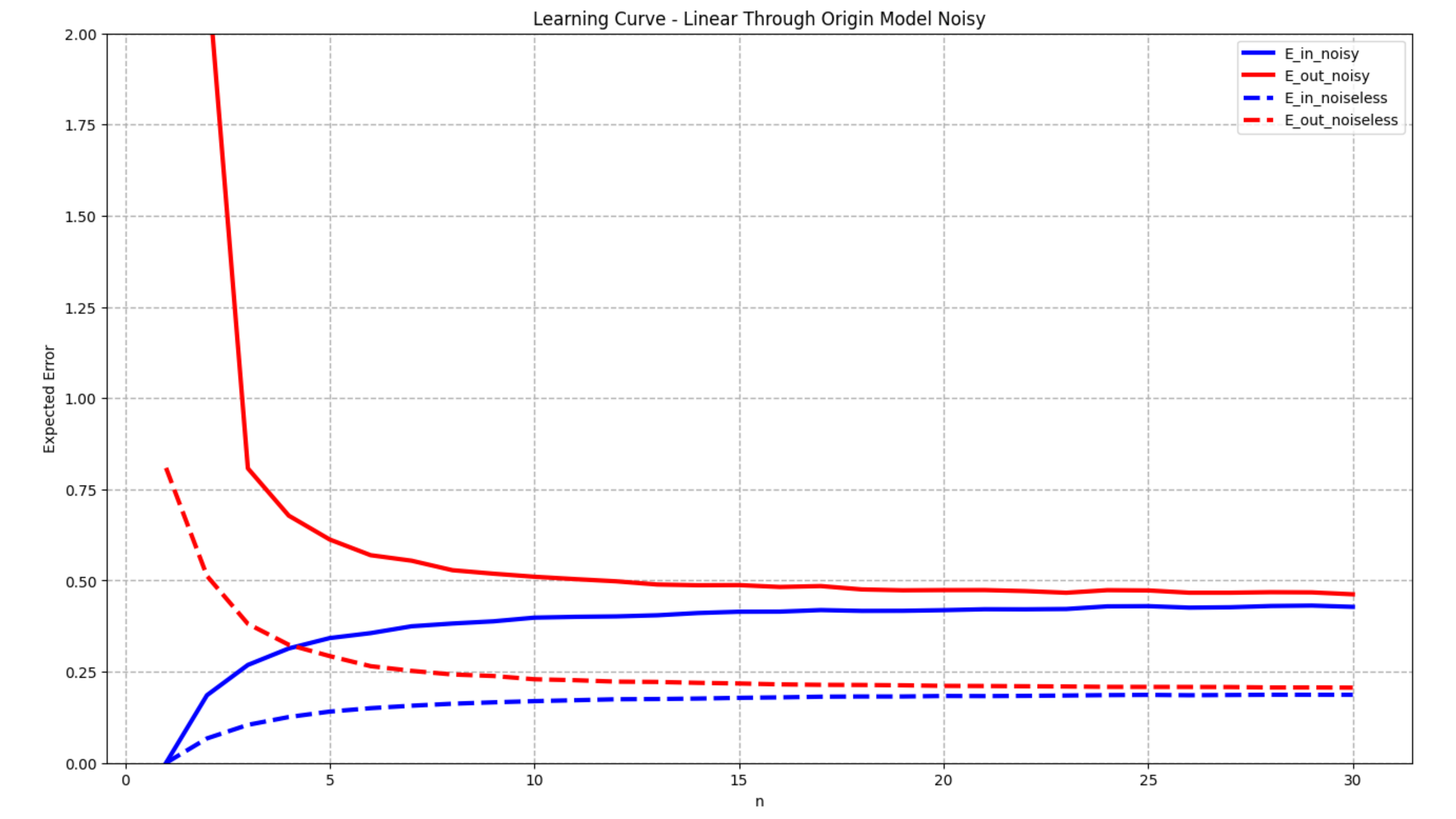

Learning Curve - Linear Model Noisy
E_in
E_out
Expected Error
n
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
5
10
15
20
25
30


Learning Curve - Linear Through Origin Model Noisy
E_in_noisy
E_out_noisy
Bias
Expected Error
n
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
0
5
10
15
20
25
30

เปรียบระหว่างไม่มีสัญญาณรบกวน กับมีสัญญาณรบกวน


Learning Curve - Constant Model Noisy
E_in_noisy
E_out_noisy
E_in_noiseless
E_out_noiseless
Expected Error
n
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
0
5
10
15
20
25
30


Learning Curve - Linear Model Noisy
E_in_noisy
E_out_noisy
E_in_noiseless
E_out_noiseless
Expected Error
n
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
5
10
15
20
25
30


Learning Curve - Linear Through Origin Model Noisy
E_in_noisy
E_out_noisy
E_in_noiseless
E_out_noiseless
Expected Error
n
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
0
5
10
15
20
25
30

Thank you