

The background image shows a stunning landscape featuring a deep blue lake curving along the base of a range of mountains. The mountains are covered in dense green forests, with some patches of snow or clouds visible at higher elevations. In the foreground, a two-lane asphalt road with white dashed lines curves along the edge of a steep cliff, bordered by a metal railing. The sky is a clear, pale blue with scattered white cumulus clouds.

DISCUSSION 6

SHUANGJIE ZHANG

NORMAL

- Symmetric
- Bell shape in 1 dimension
- Football shaped in two dimensions

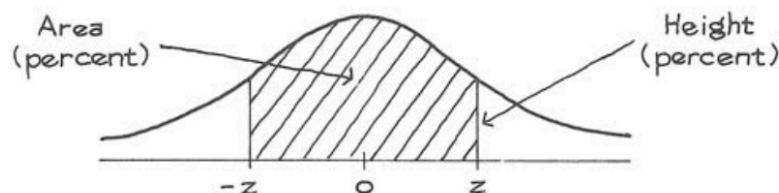
REVIEW OF PREDICTION

- 1. Visual observation
- 2. slope + intercept
- 3. z score method (standard unit)

EXERCISE

- Average SAT score = 550, SD = 80
- Average first year GPA = 2.6, SD = 0.6, $r= 0.40$
- Use z score method to compute:
- Suppose the percentile rank of one student on the SAT is 90th, among the first-year students. Predict his percentile rank on first-year GPA. The scatter plot is foot-ball shaped.

Tables



A NORMAL TABLE

<i>z</i>	Height	Area	<i>z</i>	Height	Area	<i>z</i>	Height	Area
0.00	39.89	0	1.50	12.95	86.64	3.00	0.443	99.730
0.05	39.84	3.99	1.55	12.00	87.89	3.05	0.381	99.771
0.10	39.69	7.97	1.60	11.09	89.04	3.10	0.327	99.806
0.15	39.45	11.92	1.65	10.23	90.11	3.15	0.279	99.837
0.20	39.10	15.85	1.70	9.40	91.09	3.20	0.238	99.863
0.25	38.67	19.74	1.75	8.63	91.99	3.25	0.203	99.885
0.30	38.14	23.58	1.80	7.90	92.81	3.30	0.172	99.903
0.35	37.52	27.37	1.85	7.21	93.57	3.35	0.146	99.919
0.40	36.83	31.08	1.90	6.56	94.26	3.40	0.123	99.933
0.45	36.05	34.73	1.95	5.96	94.88	3.45	0.104	99.944
0.50	35.21	38.29	2.00	5.40	95.45	3.50	0.087	99.953
0.55	34.29	41.77	2.05	4.88	95.96	3.55	0.073	99.961
0.60	33.32	45.15	2.10	4.40	96.43	3.60	0.061	99.968
0.65	32.30	48.43	2.15	3.96	96.84	3.65	0.051	99.974
0.70	31.23	51.61	2.20	3.55	97.22	3.70	0.042	99.978

0.75	30.11	54.67	2.25	3.17	97.56	3.75	0.035	99.982
0.80	28.97	57.63	2.30	2.83	97.86	3.80	0.029	99.986
0.85	27.80	60.47	2.35	2.52	98.12	3.85	0.024	99.988
0.90	26.61	63.19	2.40	2.24	98.36	3.90	0.020	99.990
0.95	25.41	65.79	2.45	1.98	98.57	3.95	0.016	99.992
1.00	24.20	68.27	2.50	1.75	98.76	4.00	0.013	99.9937
1.05	22.99	70.63	2.55	1.54	98.92	4.05	0.011	99.9949
1.10	21.79	72.87	2.60	1.36	99.07	4.10	0.009	99.9959
1.15	20.59	74.99	2.65	1.19	99.20	4.15	0.007	99.9967
1.20	19.42	76.99	2.70	1.04	99.31	4.20	0.006	99.9973
1.25	18.26	78.87	2.75	0.91	99.40	4.25	0.005	99.9979
1.30	17.14	80.64	2.80	0.79	99.49	4.30	0.004	99.9983
1.35	16.04	82.30	2.85	0.69	99.56	4.35	0.003	99.9986
1.40	14.97	83.85	2.90	0.60	99.63	4.40	0.002	99.9989
1.45	13.94	85.29	2.95	0.51	99.68	4.45	0.002	99.9991

Average SAT score = 550, SD = 80

Average first year GPA = 2.6, SD = 0.6, r= 0.40

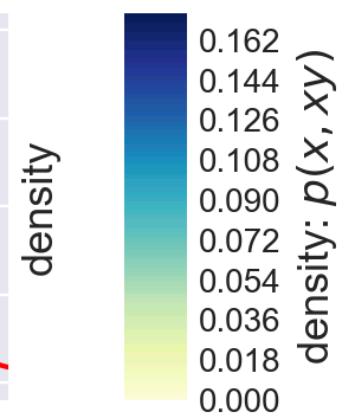
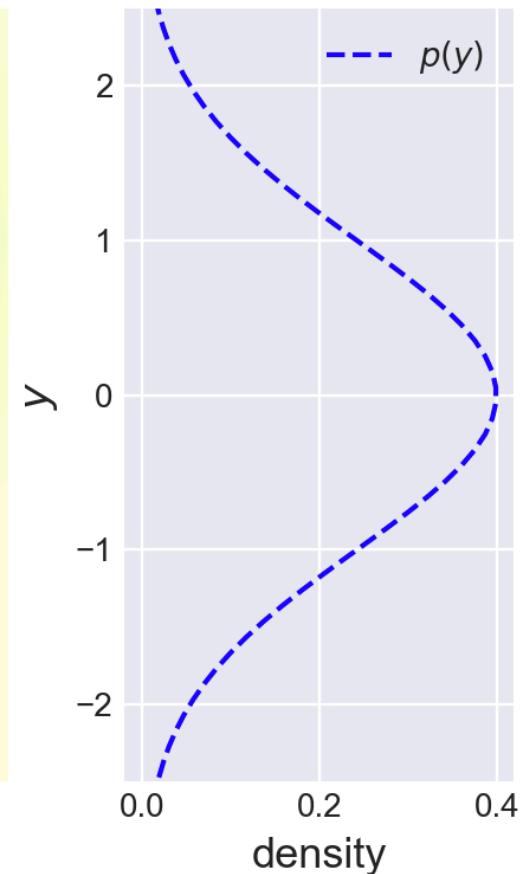
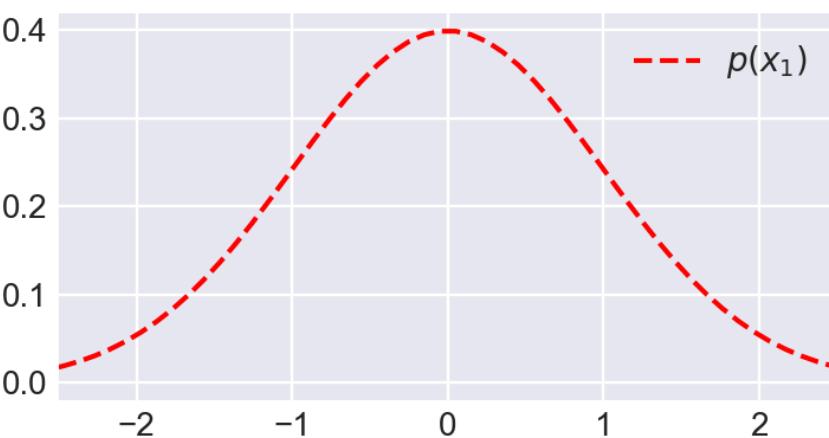
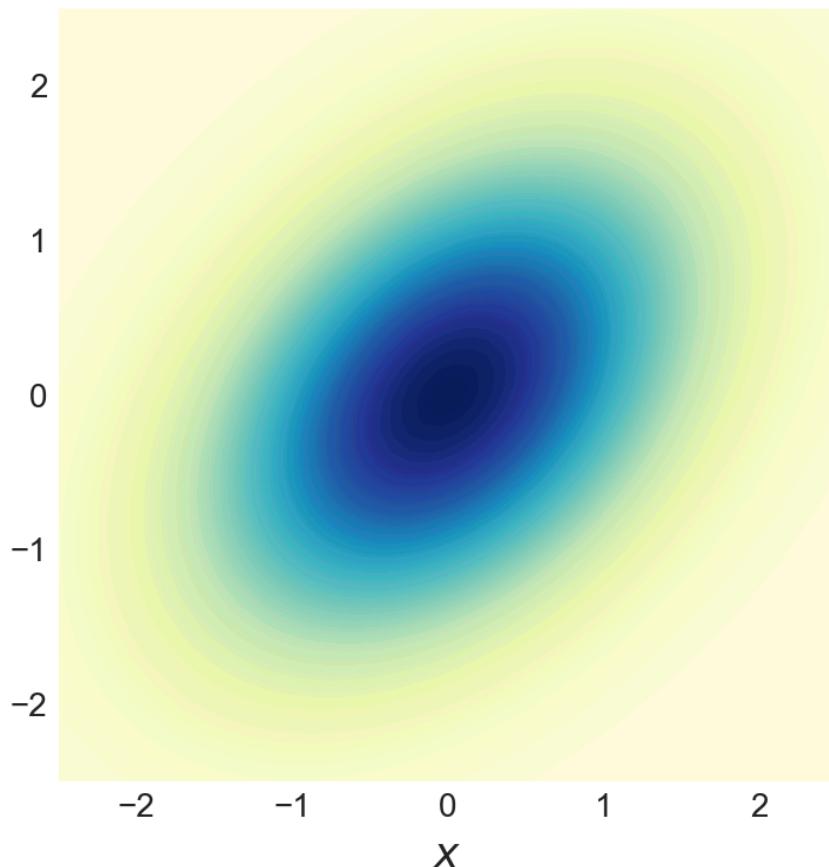
Use z score method to compute:

Suppose the percentile rank of one student on the SAT is 90th, among the first-year students. Predict his percentile rank on first-year GPA. The scatter plot is foot-ball shaped.

- Suppose the percentile rank of one student on the SAT is 90th, among the first-year students. Predict his percentile rank on first-year GPA.
- 69%

Marginal distributions

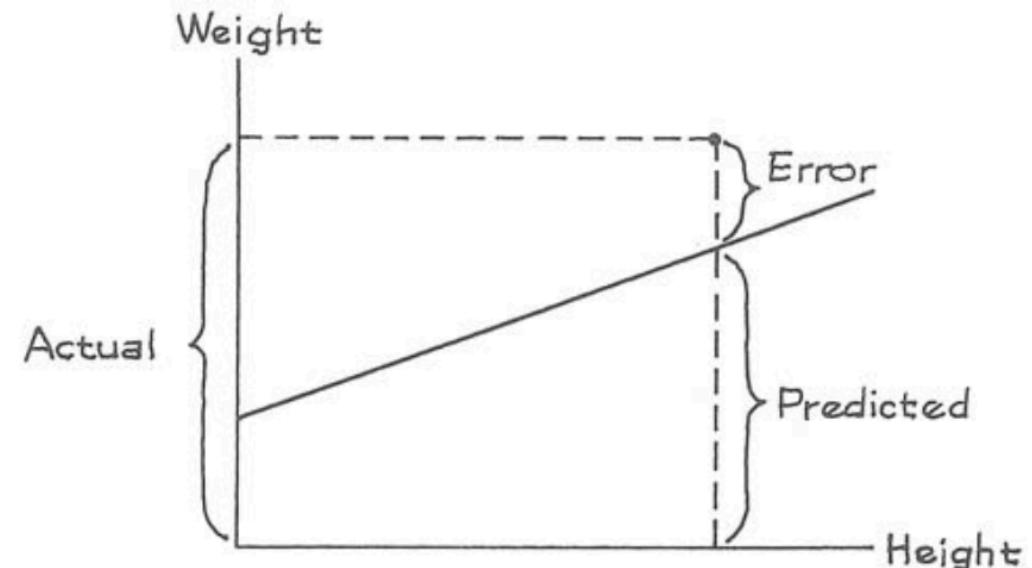
SEE FROM THE PLOT



R.M.S ERROR

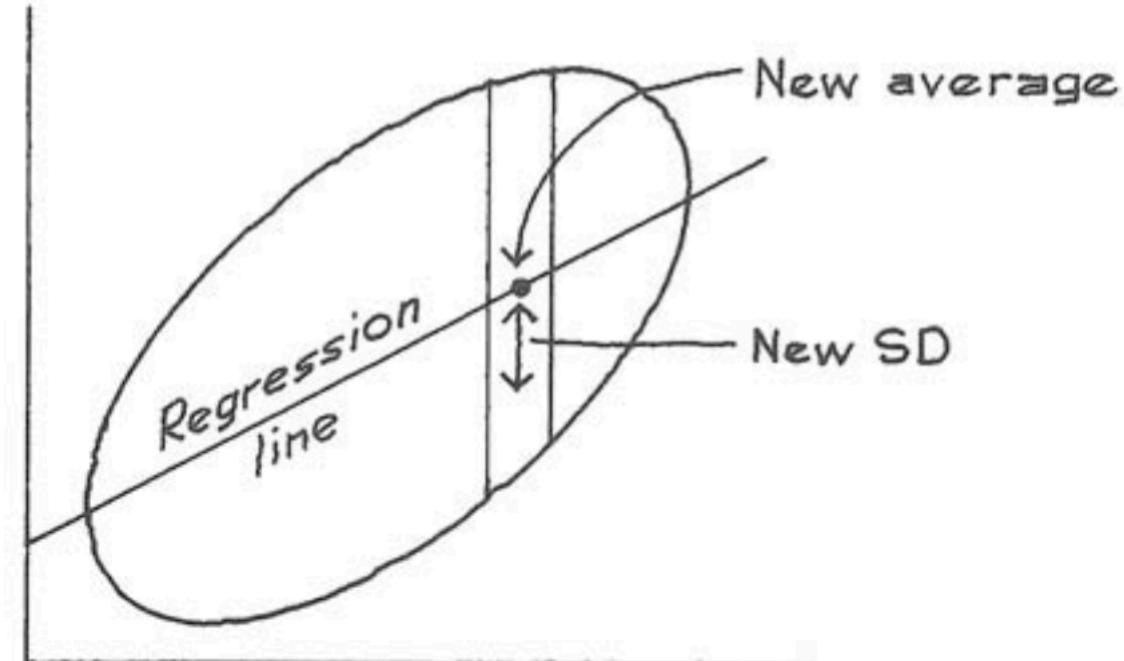
- Root Mean Square error
- R.M.S error = $\sqrt{\frac{(error \ #1)^2 + (error \ #2)^2 + \dots + (error \ #n)^2}{n}}$
- $R.M.S \ error = \sqrt{1 - r^2} \times SD \text{ of prediction variable}$
- Use actual value minus predicted value
- Square the difference
- Mean
- Square root

Figure 2. Prediction error equals vertical distance from the line.



R.M.S ERROR

- Error averages out 0.
- R.M.S Error = New SD for the prediction < SD for y



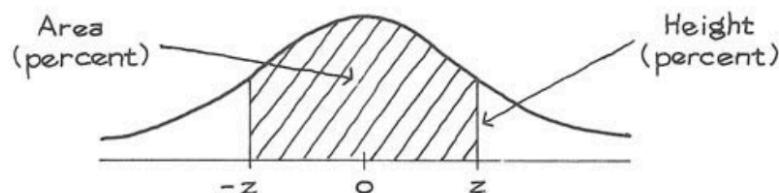
EXERCISE

- Average LSAT score = 162, SD = 6
- Average first year score = 68, SD = 10, $r= 0.60$
- Q: Among the student who scored 174 on the LSAT, about what percentage had first year scores over 86?
- Slope, intercept, RMS error
- New average
- New SD
- New z score
-

EXERCISE

- Average LSAT score = 162, SD = 6
- Average first year score = 68, SD = 10, $r= 0.60$
- Q: Among the student who scored 174 on the LSAT, about what percentage had first year scores over 88?
- Slope, intercept, RMS error: 1, -94, 8
- New average 80
- New SD 8
- New z score 0.75
- $50\% - 54.67\%/2 = 50\% - 27.335 \% = 11.3325\%$

Tables



A NORMAL TABLE

<i>z</i>	Height	Area	<i>z</i>	Height	Area	<i>z</i>	Height	Area
0.00	39.89	0	1.50	12.95	86.64	3.00	0.443	99.730
0.05	39.84	3.99	1.55	12.00	87.89	3.05	0.381	99.771
0.10	39.69	7.97	1.60	11.09	89.04	3.10	0.327	99.806
0.15	39.45	11.92	1.65	10.23	90.11	3.15	0.279	99.837
0.20	39.10	15.85	1.70	9.40	91.09	3.20	0.238	99.863
0.25	38.67	19.74	1.75	8.63	91.99	3.25	0.203	99.885
0.30	38.14	23.58	1.80	7.90	92.81	3.30	0.172	99.903
0.35	37.52	27.37	1.85	7.21	93.57	3.35	0.146	99.919
0.40	36.83	31.08	1.90	6.56	94.26	3.40	0.123	99.933
0.45	36.05	34.73	1.95	5.96	94.88	3.45	0.104	99.944
0.50	35.21	38.29	2.00	5.40	95.45	3.50	0.087	99.953
0.55	34.29	41.77	2.05	4.88	95.96	3.55	0.073	99.961
0.60	33.32	45.15	2.10	4.40	96.43	3.60	0.061	99.968
0.65	32.30	48.43	2.15	3.96	96.84	3.65	0.051	99.974
0.70	31.23	51.61	2.20	3.55	97.22	3.70	0.042	99.978

0.75	30.11	54.67	2.25	3.17	97.56	3.75	0.035	99.982
0.80	28.97	57.63	2.30	2.83	97.86	3.80	0.029	99.986
0.85	27.80	60.47	2.35	2.52	98.12	3.85	0.024	99.988
0.90	26.61	63.19	2.40	2.24	98.36	3.90	0.020	99.990
0.95	25.41	65.79	2.45	1.98	98.57	3.95	0.016	99.992
1.00	24.20	68.27	2.50	1.75	98.76	4.00	0.013	99.9937
1.05	22.99	70.63	2.55	1.54	98.92	4.05	0.011	99.9949
1.10	21.79	72.87	2.60	1.36	99.07	4.10	0.009	99.9959
1.15	20.59	74.99	2.65	1.19	99.20	4.15	0.007	99.9967
1.20	19.42	76.99	2.70	1.04	99.31	4.20	0.006	99.9973
1.25	18.26	78.87	2.75	0.91	99.40	4.25	0.005	99.9979
1.30	17.14	80.64	2.80	0.79	99.49	4.30	0.004	99.9983
1.35	16.04	82.30	2.85	0.69	99.56	4.35	0.003	99.9986
1.40	14.97	83.85	2.90	0.60	99.63	4.40	0.002	99.9989
1.45	13.94	85.29	2.95	0.51	99.68	4.45	0.002	99.9991

Quiz:

Average SAT score = 550, SD = 80

Average first year GPA = 2.6, SD = 0.6, r= 0.40

Suppose the percentile rank of one student on the SAT is 80th , among the first-year students. Predict his percentile rank on first-year GPA. The scatter plot is football shaped.