

## Anomaly Detection

- Suppose you are developing an anomaly detection system to catch manufacturing defects in airplane engines.
- Your model uses

$$p(x) = \prod_{j=1}^n p(x_j; \mu_j, \sigma_j^2)$$

- You have two features  $x_1 = \textit{vibration intensity}$ , and  $x_2 = \textit{heat generated}$ .
- Both  $x_1$  and  $x_2$  take on values between 0 and 1 (and are strictly greater than 0), and for most "normal" engines you expect that  $x_1 \approx x_2$ .
- One of the suspected anomalies is that a flawed engine may vibrate very intensely even without generating much heat (large  $x_1$ , small  $x_2$ ), even though the particular values of  $x_1$  and  $x_2$  may not fall outside their typical ranges of values.
- What additional feature  $x_3$  should you create to capture these types of anomalies:

### Solution Options

- $x_3 = x_1 + x_2$  This could take on large or small values for both normal and anomalous examples, so it is not a good feature.