# EMSE 6035: Marketing of Technology

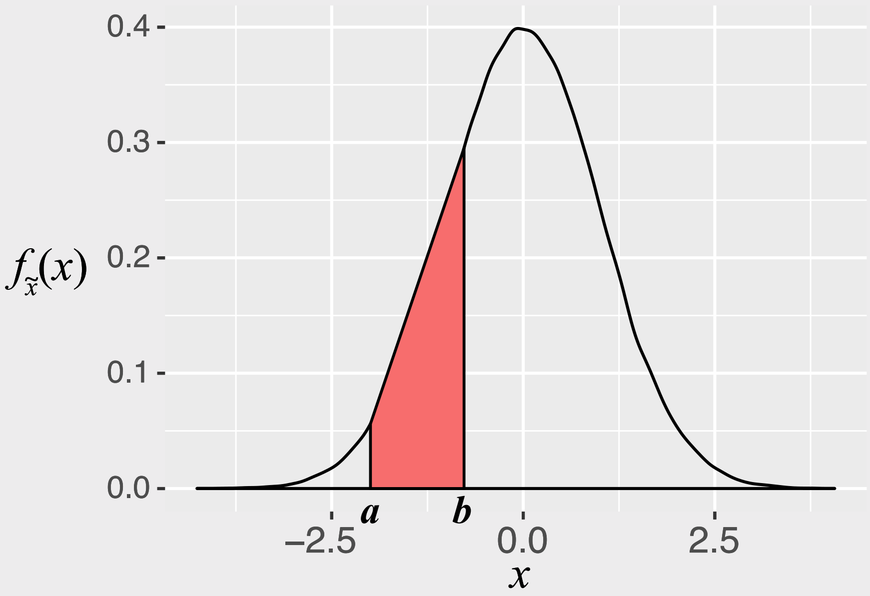
## Intro to choice modeling

John Paul Helveston, Ph.D.
Assistant Professor
Engineering Management & Systems Engineering
The George Washington University

# Review of Probability Concepts

Random Variable, $\tilde{x}$: A variable whose value is subject to variation due to chance

**Probability Density Function (PDF)**

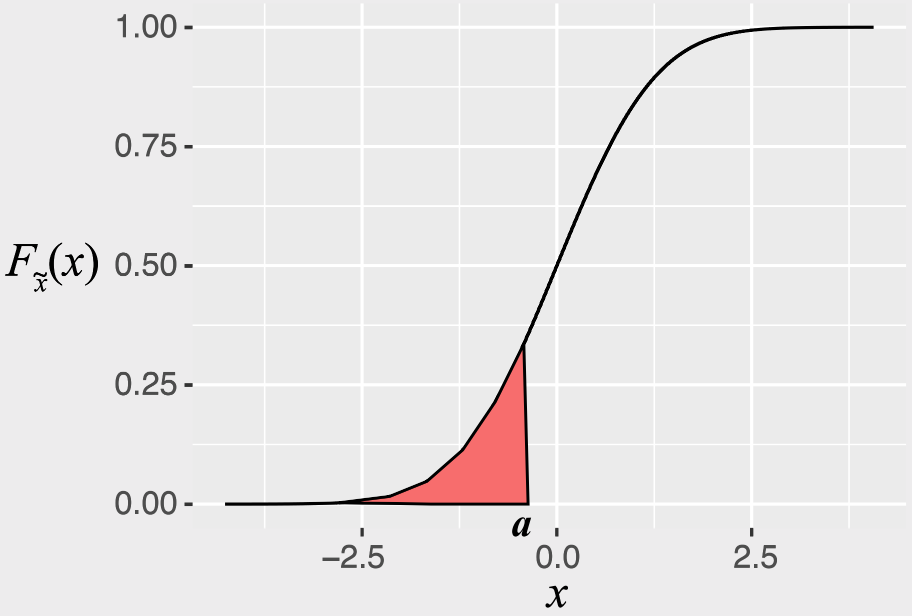



Ex: The probability that the value of $\tilde{x}$ is between $a$ and $b$ is:

$$\Pr(a \le \tilde{x} \le b) = \int_a^b f_{\tilde{x}}(x)\, dx$$

PDF

**Cumulative Density Function (CDF)**



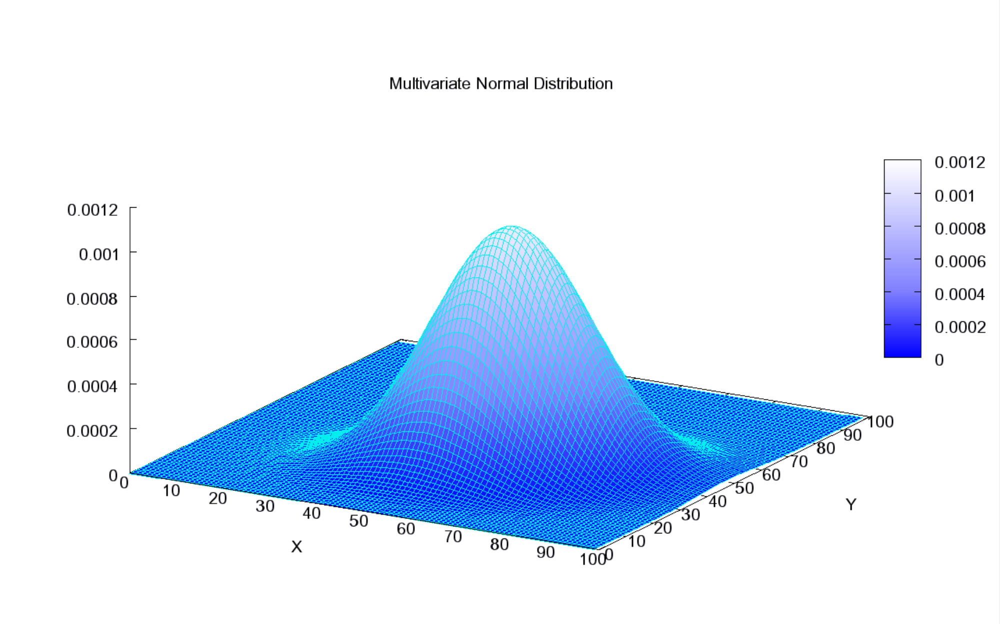

Ex: The probability that the value of $\tilde{x}$ is less than or equal to value $a$ is:

$$F_{\tilde{x}}(a) = \Pr(\tilde{x} \le a) = \int_{-\infty}^{a} f_{\tilde{x}}(x)\, dx$$

# Review of Probability Concepts

## Multivariable Joint Distribution

Joint distribution of multiple random variables, e.g. $\tilde{x}$ and $\tilde{y}$ in this figure:




$$f_{\tilde{x},\tilde{y}}(x, y)$$

https://en.wikipedia.org/wiki/Multivariate_normal_distribution

## Independence

Random variables are "independent" if the value of one doesn't affect the other's probability

In this case, the joint probability distribution is the multiplication of the individual distributions:

$$f_{\tilde{x},\tilde{y}}(x, y) = f_{\tilde{x}}(x) f_{\tilde{y}}(y)$$

# Practice Question 1

a) A random variable, $\tilde{x}$, has the PDF, $f_{\tilde{x}}(x)$. Write the equation to compute its total probability (hint: think area under the curve!). What is the answer to the equation?

b) A random variable, $\tilde{x}$, has a uniform distribution between the values 0 and 1. Draw the probability density function (PDF) and Cumulative Density Function (CDF) of $\tilde{x}$.

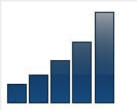
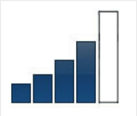
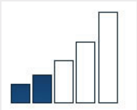
c) The value of a random variable, $\tilde{x}$, is determined by rolling one fair, 6-sided dice. Draw the PDF and CDF of $\tilde{x}$.

# Random Utility Theory

<u>Utility</u>: The satisfaction a consumer receives from a product

- Utility is a random variable (so we give it a squiggly line hat): $\tilde{u}_j$
- Utility has *relative* (not absolute) value
- Utility is unit-less

Example:

| **Attribute** | **Phone 1** | **Phone 2** | **Phone 3** |
|---|---|---|---|
| Price | \$200 | \$300 | \$400 |
| Battery Life |  |  |  |
| Signal Quality |  |  |  |
| *Utility* | $\tilde{u}_1$ > | $\tilde{u}_2$ > | $\tilde{u}_3$ |

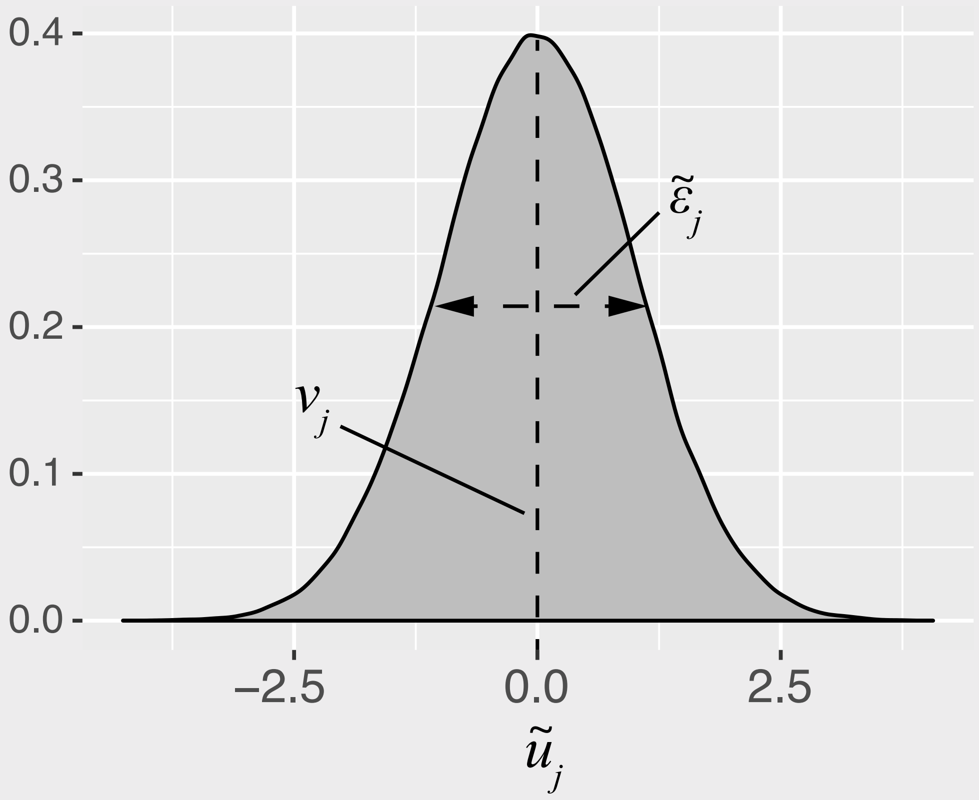
# Random Utility Theory

Utility can be broken into two parts:

$$\tilde{u}_j = v_j + \tilde{\varepsilon}_j$$

Things we can observe / measure ("Observed Utility")

Things we can't observe / measure ("Error")

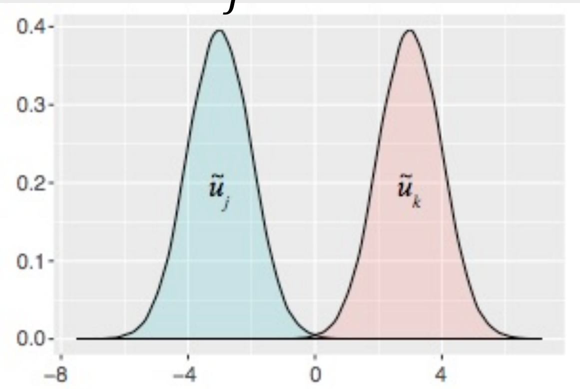


# Random Utility Theory

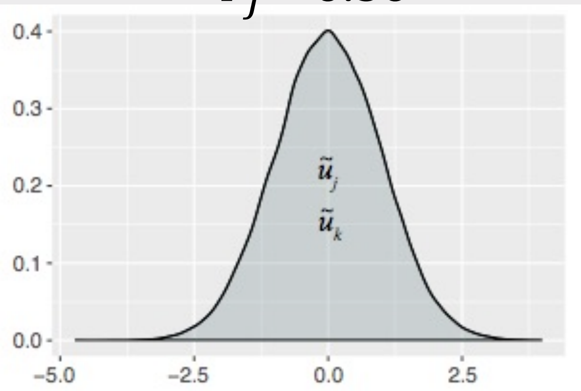
We assume that a consumer will choose product $j$ over $k$ if: $\tilde{u}_j > \tilde{u}_k$

Since utility is a *random variable*,
we can only compute the *probability* that $\tilde{u}_j > \tilde{u}_k$: $P_j = \Pr(\tilde{u}_j > \tilde{u}_k)$

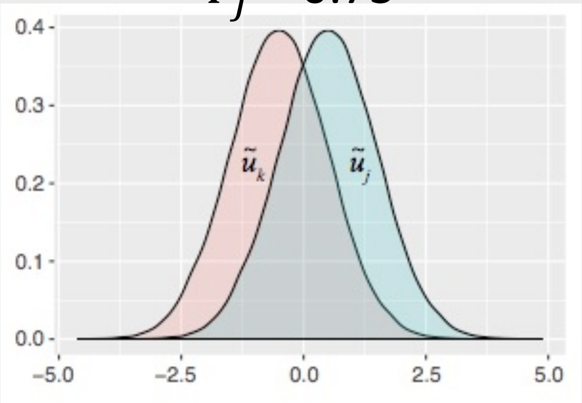
$P_j$ = 0.001



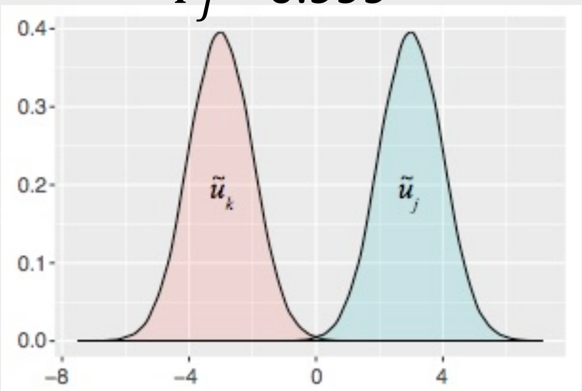

$P_j$ = 0.50




$P_j$ = 0.75



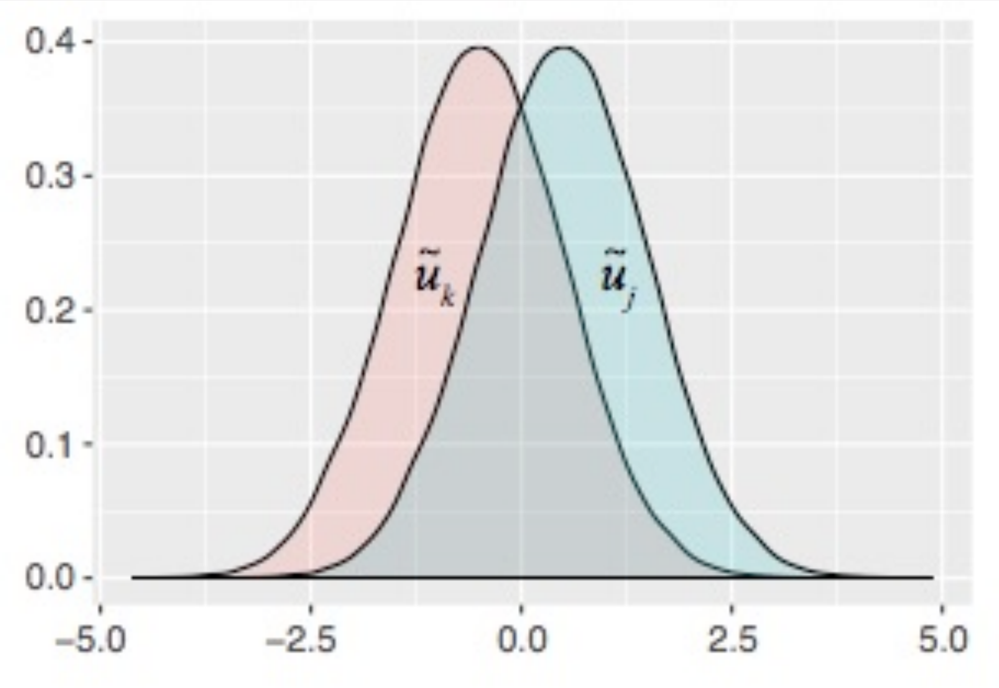

$P_j$ = 0.999



# Random Utility Theory

To compute the probability that the consumer will choose product $j$ over $k$, we have to integrate over the joint distribution:

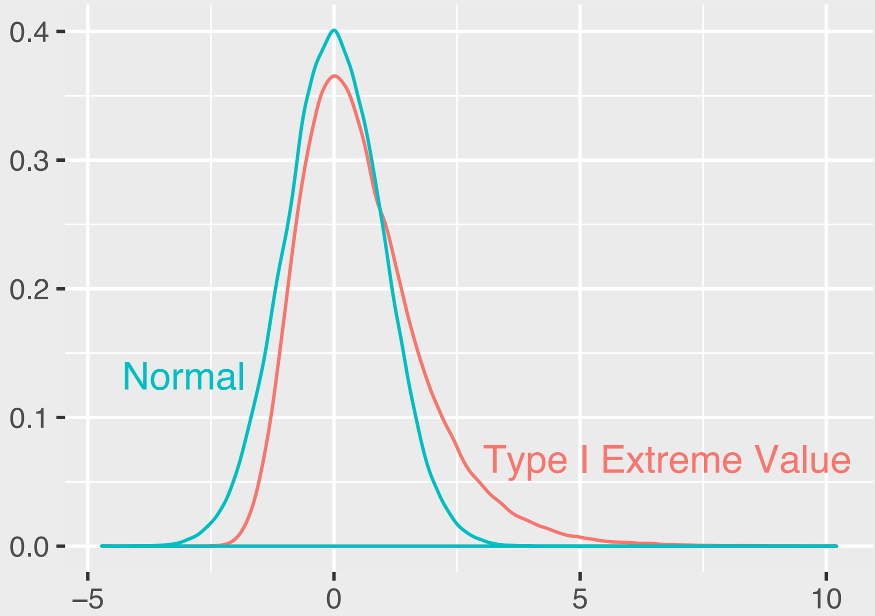
$$
\begin{aligned}
P_j &= \Pr(\tilde{u}_j > \tilde{u}_k) \\
&= \Pr(v_j + \tilde{\varepsilon}_j > v_k + \tilde{\varepsilon}_k) \\
&= \Pr(\tilde{\varepsilon}_k \leq \tilde{\varepsilon}_j + (v_j - v_k)) \\
&= \int_{\varepsilon_j=-\infty}^{\infty} \left[ \int_{\varepsilon_k=-\infty}^{\tilde{\varepsilon}_j+(v_j-v_k)} f_{\tilde{\varepsilon}}\left(\varepsilon_j, \varepsilon_k\right) d\varepsilon_k \right] d\varepsilon_j
\end{aligned}
$$




To solve this equation, we need to assume a distribution for $\tilde{\varepsilon}$

# Assumptions for $\tilde{\varepsilon}$

| Name | Distribution | Advantage | Disadvantage |
|---|---|---|---|
| Probit | Normal<br>$\tilde{\varepsilon} \sim N(0, \Sigma)$ | Supported by theory | No closed for solution for integral |
| Logit | Type I Extreme Value | Closed form solution for probabilities! | Strong assumptions:<br>- Errors must be independent |




$$P_j = \frac{e^{v_j}}{\sum_{k=1}^{J} e^{v_k}}$$

<u>Ex</u>: Probability of choosing *j vs. k*:

$$P_j = \frac{e^{v_j}}{e^{v_j} + e^{v_k}}, \quad P_k = \frac{e^{v_k}}{e^{v_j} + e^{v_k}}$$

(we'll be using the <u>logit</u> model for this class!)

# Independence of Irrelevant Alternatives (IIA)

Logit model has the IIA property, which can be problematic when products are close substitutes

Classic Example: “Red Bus” vs. “Blue Bus”

| | Taxi | Red Bus |
|---|---|---|
| $e^{v_j}$ | 2 | 1 |
| $P_j$ | $\frac{2}{2+1}$ $= 0.66$ | $\frac{1}{2+1}$ $= 0.33$ |

$$P_j = \frac{e^{v_j}}{\sum_{k=1}^{J} e^{v_k}}$$

| | Taxi | Red Bus | Blue Bus |
|---|---|---|---|
| $e^{v_j}$ | 2 | 1 | 1 |
| $P_j$ | $\frac{2}{2+1+1}$ $= 0.50$ | $\frac{1}{2+1+1}$ $= 0.25$ | $\frac{1}{2+1+1}$ $= 0.25$ |
| We expect the probabilities to be: | $0.66$ | $0.165$ | $0.165$ |

# Practice Question 2

a) A consumer is making a choice between two bars of chocolate: milk chocolate (*m*) and dark chocolate (*d*). Assume that we know the observed utility of each bar to be $v_m = 3$ and $v_d = 4$. Using a logit model, compute the probabilities of choosing each bar: $P_m$ and $P_d$.

Hint:

$$P_j = \frac{e^{v_j}}{\sum_{k=1}^{J} e^{v_k}}$$

b) A third bar of chocolate is now added to the choice set. It is the exact same as the milk chocolate bar, but it has a slightly different wrapper (which has no effect on the consumer's utility). Now, $v_{m1} = v_{m2} = 3$, and $v_d = 4$. Based on the probabilities from question 2a, what would we *expect* the probabilities of choosing each bar to be? What probabilities does the logit model produce?

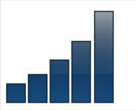
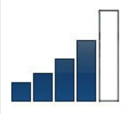
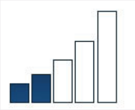
# How do we get $v_j$?

We define $v_j$ as a function of observable product attributes, $x_j$:

$$v_j = f(x_j) = \beta_1 x_{j1} + \beta_2 x_{j2} + \ldots$$

Weights that denote the *relative* value of attributes $x_{j1}$ and $x_{j2}$

Example:

| **Attribute** | **Phone 1** | **Phone 2** | **Phone 3** |
|---|---|---|---|
| Price | \$200 | \$300 | \$400 |
| Battery Life |  |  |  |
| Signal Quality |  |  |  |

# How do we get $v_j$?

We define $v_j$ as a function of observable product attributes, $x_j$:

$$v_j = f(x_j) = \beta_1 x_{j1} + \beta_2 x_{j2} + \ldots$$

Weights that denote the *relative* value of attributes $x_{j1}$ and $x_{j2}$

Example:

| | Attribute | Phone 1 | Phone 2 | Phone 3 |
|---|---|---|---|---|
| $x_1$ | Price | $200 | $300 | $400 |
| $x_2$ | Battery Life (hours) | 20 | 15 | 10 |
| $x_3$ | Signal Quality | 100% | 80% | 60% |

$x_1$ $x_2$ $x_3$

Phone 1: $v_1 = \beta_1(200) + \beta_2(20) + \beta_3(100)$

Phone 2: $v_2 = \beta_1(300) + \beta_2(15) + \beta_3(80)$

Phone 3: $v_3 = \beta_1(400) + \beta_2(10) + \beta_3(60)$

Example: Let's say $\beta_1 = -0.01$, $\beta_2 = 0.1$, $\beta_3 = 0.05$

Phone 1: $v_1 = -0.01(200) + 0.01(20) + 0.02(100) = 5$

Phone 2: $v_1 = -0.01(300) + 0.01(15) + 0.02(80) = 2.5$

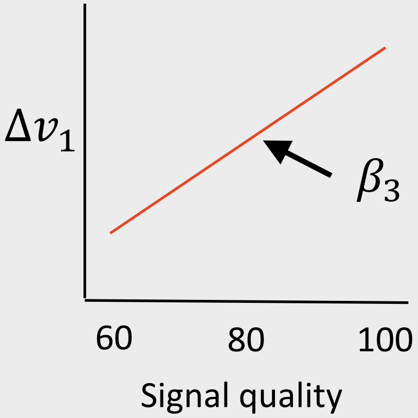
Phone 3: $v_1 = -0.01(400) + 0.01(10) + 0.02(60) = 0$

# Continuous vs. Discrete Attributes ($x_j$)

## Continuous attributes

| | Attribute | Phone 1 | Phone 2 | Phone 3 |
|---|---|---|---|---|
| $x_1$ | Price | $200 | $300 | $400 |
| $x_2$ | Battery Life (hours) | 20 | 15 | 10 |
| $x_3$ | Signal Quality | 100% | 80% | 60% |

Phone 1: $v_1 = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$



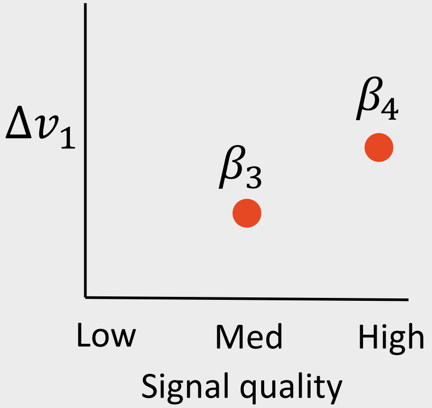

## Discrete (categorical) attributes

| Attribute | Phone 1 | Phone 2 | Phone 3 |
|---|---|---|---|
| Price | $200 | $300 | $400 |
| Battery Life (hours) | 20 | 15 | 10 |
| Signal Quality | High | Med | Low |

Phone 1: $v_1 = \beta_1 x_1 + \beta_2 x_2 + \beta_3 \delta^{\text{MED}} + \beta_4 \delta^{\text{HIGH}}$

$\delta^{\text{MED}} = 1 \; or \; 0$

$\delta^{\text{HIGH}} = 1 \; or \; 0$



# Practice Question 3

| Attribute | Bar 1 | Bar 2 | Bar 3 |
|---|---|---|---|
| Price | $1.20 | $1.50 | $3.00 |
| % Cacao | 10% | 60% | 80% |

a) Write out a model for the observed utility of each chocolate bar in the above set.

b) If the coefficient for the *price* attribute was -0.1 and the coefficient for *% Cacao* attribute was 0.1, what is the difference in the observed utility between bars 3 and 1?

c) With the addition of the *brand* attribute, repeat part *a*.

# Extra Slides

Let's say our utility function is:

$$u_j = \beta_1 x_j^{price} + \beta_2 x_j^{cacao} + \beta_3 \delta_j^{\text{hersheys}} + \beta_4 \delta_j^{\text{lindt}} + \varepsilon_j$$

And we estimate the following coefficients:

| Parameter | Coef. |
|---|---|
| $\beta_1$ | -0.1 |
| $\beta_2$ | 0.1 |
| $\beta_3$ | -2.0 |
| $\beta_4$ | -0.1 |

a) What are the expected probabilities of choosing each bar using a logit model?

| **Attribute** | **Bar 1** | **Bar 2** | **Bar 3** |
|---|---|---|---|
| Price | \$1.20 | \$1.50 | \$3.00 |
| % Cacao | 10% | 60% | 80% |
| Brand | Hersheys | Lindt | Ghirardelli |

b) What price would Bar 2 have to be to get a 50% market share?