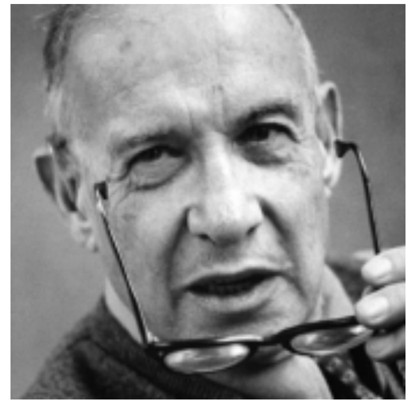
Choice Modeling

Marketing Analytics

Professor Kamel Jedidi
Columbia University

The aim of marketing is to *know and understand* the customer so well the product or service fits him and sells itself.



Peter Drucker

Founder of Modern Management

Consumer Choice Behavior

- Consumers decide on:
  - Whether to buy (purchase incidence)
  - What to buy (brand consideration and choice)
  - Where to buy (channel choice)
  - Whether to buy again (loyalty/churn)

- Marketers learn about consumers through:
  - What they do (revealed choice)
  - What they say (stated choice)

Uses of Choice Modeling in Marketing

- Segmentation and targeting
- Product positioning
- Marketing mix decisions
  - Product design
  - Pricing
  - Advertising and promotions
  - Customer churn management

Example: Product Design
Choice-Based-Conjoint (CBC)

Brand	Hilton	WESTIN	NEW HOTEL	None – I would not choose any of these.
Location	10 minute ride to destination	30 minute ride to destination	Walking distance to destination	
Restaurant	Restaurant within walking distance	Restaurant within 5 minute car ride	Restaurant in hotel	
Gym	No gym	On-site gym	Partner gym within 5 minutes	
Wireless	Wireless Internet connection throughout the hotel	Wireless “hot spots” in the hotel, but not in the room	No wireless access	
Rewards	Earn Standard Rewards Points	Earn Double Rewards Points	Earn Triple Rewards Points	
Room Rate	$200	$225	$150	
Choose one	A	B	C	

Example: Scanner-Panel Data
Nielsen HomeScan

By scanning the items you purchase (from cereal at the store to a candy bar in a snack machine) retailers see where you shop, what you buy, …

Sample Scanner-Panel Data

Customer ID	Date	Store ID	Brand	Quantity	Regular Price	Discount	Display	Feature
1001	3/1/2016	2345	Tide	50oz	$3.55	$0.43	No	No
1001	3/29/2016	5678	Tide	64oz	$3.99	$0.54	Yes	Yes
1001	4/25/2016	2345	Tide	50oz	$3.55	$0.45	No	No
1001	5/28/2016	5678	All	50oz	$2.99	$0.50	Yes	No
1001	6/27/2016	2345	Tide	50oz	$3.60	$0.45	No	No
1001	7/22/2016	5678	Tide	50oz	$3.60	$0.20	No	No
1001	8/29/2016	2345	All	64oz	$3.15	$0.60	Yes	Yes
1001	9/24/2016	5678	Tide	50oz	$3.65	$0.42	No	No
1001	10/28/2016	2345	All	50oz	$4.99	$1.00	Yes	Yes
1001	11/25/2016	5678	Tide	50oz	$3.99	$0.50	No	No

Example: Monthly Churn Rate for Wireless Carriers in U.S.

- Monthly churn rates in 2017 (Q2)
  - Verizon 1.19%
  - AT&T 1.28%
  - U.S. Cellular 1.52%
  - Sprint 2.24%
  - T-Mobile 2.28%
  - Shentel 2.88%

  - **Average 1.90%**

Source: https://www.statista.com/statistics/283511/average-monthly-churn-rate-top-wireless-carriers-us/

Outline

- Logistic regression (binary choice)
- Multinomial logit (multiple choice)

Logistic Regression Motivation

- Online/Catalog purchase (Buy/No-Buy)
  - Recency, Frequency, Monetary value (RFM) measures as predictors of purchase
- Response to marketing efforts
  - Did the customer buy after being sent a coupon or an email ad?
- Churn
  - Can we predict customer churn before it happens?

What is Common to these Examples?

- The outcome variable is binary
  - Coded: Y = 1 (if “Yes”) and Y = 0 (if “No”)

- There is a set of variables (x’s) that we can use to explain and predict the binary outcome variable

Example - Catalog Data

- Explanatory Variables
  - Recency – how many days since last purchase
  - Frequency – how many times the consumer buys
  - Monetary Value – Total $ amount spent

- Dependent Variable
  - Purchase (Yes/No)

Excerpt from the RFM Data

```
RFMdata <- read.csv(file = "RFMData.csv",row.names=1)
kable(head(RFMdata,5),row.names = TRUE)
```

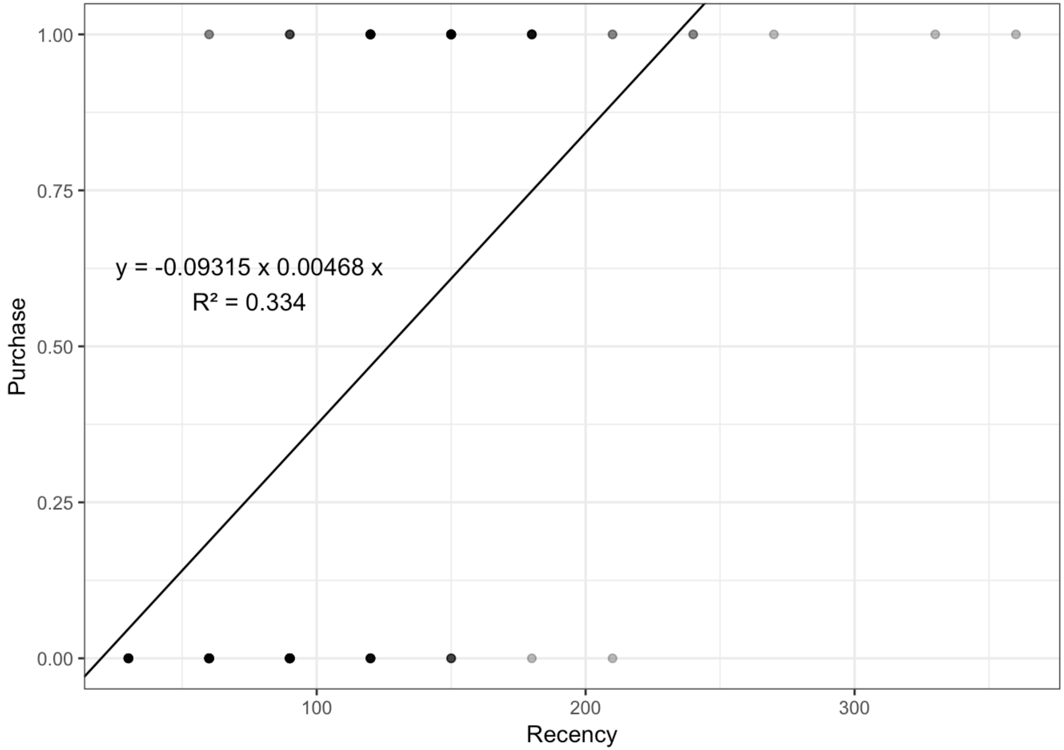
	Recency	Frequency	Monetary	Purchase
1	120	7	41.66	0
2	90	9	46.71	0
3	120	6	103.99	1
4	270	17	37.13	1
5	60	5	88.92	0

Purchase rate in RFM data=45/100=45%

Why Can’t We Just Use Regression?

- Predictions could be outside the range of [0,1] interval
- Statistical tests from regression would be wrong

Recency vs. Purchase



Logistic Regression Model

- The model states that a consumer has a utility (a desire) from buying and a utility from not buying (keep the money)
- Utility from buying: V_b
- Utility from not buying: $V_n=0$
- Consumer buys if $V_b > V_n=0$

The Choice Probability

The probability of buying is proportional to its utility (i.e., attractiveness):

$$p = \frac{\exp(V_b)}{\exp(V_b) + \exp(V_n)} = \frac{\exp(V_b)}{\exp(V_b) + 1}$$

Example

- Utility from buying: $V_b = 2$
- Utility from not buying: $V_n = 0$
- Probability of buying:

$$p = \frac{\exp(2)}{\exp(2)+1} = \frac{7.39}{7.39+1} = 0.88$$

- Odds of buying

$$\frac{p}{1-p} = \frac{0.88}{1-0.88} = 7.39 = \exp(2)$$

Utility Varies across Customers as a Function of RFM variables

- For RFM data, the utility of buying:

$V_b = \beta_0 + \beta_1 \text{Recency} + \beta_2 \text{Frequency} + \beta_3 \text{Monetary}$

- Logistic regression software uses the data to estimate the model parameters (the betas)

R-Code for Logistic Regression

Outcome Variable

Independent Variables

Binary Outcome

```
model <- glm(Purchase~Recency+Frequency+Monetary, data=RFMdata, family = "binomial")
output <- cbind(coef(summary(model))[, 1:4],exp(coef(model)))
colnames(output) <- c("beta","SE","z val.","Pr(>|z|)",'exp(beta)')
kable(output,caption = "Logistic regression estimates")
```

See Logistic Regression R Notebook for programming details.

Logistic Regression Output

```r
# likelihood ratio test
reduced.model <- glm(Purchase ~ 1, data=RFMdata, family = "binomial")
kable(xtable(anova(reduced.model, model, test = "Chisq")),caption = "Likelihood ratio test")
```

Likelihood ratio test

Resid. Df	Resid. Dev	Df	Deviance	Pr(>Chi)
99	137.62776	NA	NA	NA
96	30.48715	3	107.1406	0

Observed χ^2

P-value

Significance-level

Likelihood ratio test:

Assess overall significance

Logistic Regression Output

Logistic regression estimates

	beta	SE	z val.	Pr(>|z|)	exp(beta)
(Intercept)	-30.2976692	8.5522913	-3.542638	0.0003961	0.000000
Recency	0.1114175	0.0309797	3.596464	0.0003226	1.117862
Frequency	0.5941268	0.2429393	2.445577	0.0144620	1.811448
Monetary	0.1677054	0.0465645	3.601572	0.0003163	1.182588

Regression coefficients measure impact of x on utility

t-test for significance

How Do You Interpret Exp(beta)?

Logistic regression estimates

	beta	SE	z val.	Pr(>|z|)	exp(beta)
(Intercept)	-30.2976692	8.5522913	-3.542638	0.0003961	0.000000
Recency	0.1114175	0.0309797	3.596464	0.0003226	1.117862
Frequency	0.5941268	0.2429393	2.445577	0.0144620	1.811448
Monetary	0.1677054	0.0465645	3.601572	0.0003163	1.182588

Interpretation of Exp(beta)

- Consider two consumers (1 & 2) with identical values on recency and frequency, but consumer 1 spends $1 more than consumer 2.
  - Then the odds of buying for consumer 1 are 1.183 the odds of consumer 2.
  - More generally, the odds of buying are 18.3% higher for each increase of Monetary Value by $1.

Predicting Purchase Probabilities

- Estimated utility function in RFM data*:*

$$V= -30.29 + .111\text{Recency} + .594\text{Frequency} + .168\text{Monetary}$$

- Logistic regression predictions

$$p = \frac{\exp(V)}{\exp(V)+1}$$

Predicting Purchase Probabilities in R

```
# calculate logit probabilities
RFMdata$Base.Probability <- predict(model, RFMdata, type="response")
kable(head(RFMdata,5),row.names = TRUE)
```

	Recency	Frequency	Monetary	Purchase	Probability
1	120	7	41.66	0	0.0030728
2	90	9	46.71	0	0.0008332
3	120	6	103.99	1	0.9833225
4	270	17	37.13	1	0.9999999
5	60	5	88.92	0	0.0032378

Classification

All people with probability less ½ ➔ No purchase
All people with probability above ½ ➔ Purchase

```
# purchase vs. no purchase <-> p>0.5 or p<0.5
RFMdata$Predicted.Purchase <- 1*(RFMdata$Base.Probability>=0.5)
kable(head(RFMdata,5),row.names = TRUE)
```

	Recency	Frequency	Monetary	Purchase	Base.Probability	Predicted.Purchase
1	120	7	41.66	0	0.0030728	0
2	90	9	46.71	0	0.0008332	0
3	120	6	103.99	1	0.9833225	1
4	270	17	37.13	1	0.9999999	1
5	60	5	88.92	0	0.0032378	0

Classification (Hit Rate)

```
Confusion Matrix and Statistics

          Reference
Prediction  0  1
 No Buy 0  51  2
 Buy    1   4 43

               Accuracy : 0.94
                 95% CI : (0.874, 0.9777)
    No Information Rate : 0.55
    P-Value [Acc > NIR] : <2e-16

                  Kappa : 0.8793
 Mcnemar's Test P-Value : 0.6831

            Sensitivity : 0.9556
            Specificity : 0.9273
         Pos Pred Value : 0.9149
         Neg Pred Value : 0.9623
             Prevalence : 0.4500
         Detection Rate : 0.4300
   Detection Prevalence : 0.4700
      Balanced Accuracy : 0.9414

       'Positive' Class : 1
```

```
confusionMatrix(RFMdata$Predicted.Purchase,RFMdata$Purchase,positive = "1")
```

Hit Rate=(51+43)/100=94%

True positive rate (Recall)=43/(43+2)=96%

True negative rate=51/(51+4)=93%

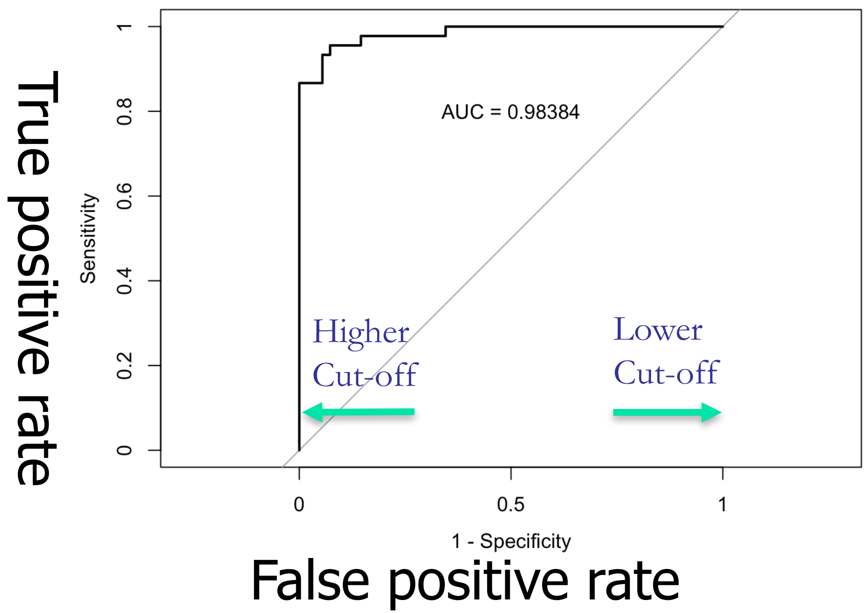
False positive rate =1-93%=7%

ROC (Receiver Operating Characteristic) Curve

```r
library(pROC)
rocobj <- roc(RFMdata$Purchase, RFMdata$Base.Probability)
{plot(rocobj,legacy.axes=TRUE)
text(0.5, 0.8, labels = sprintf("AUC = %.5f",rocobj$auc))}
```

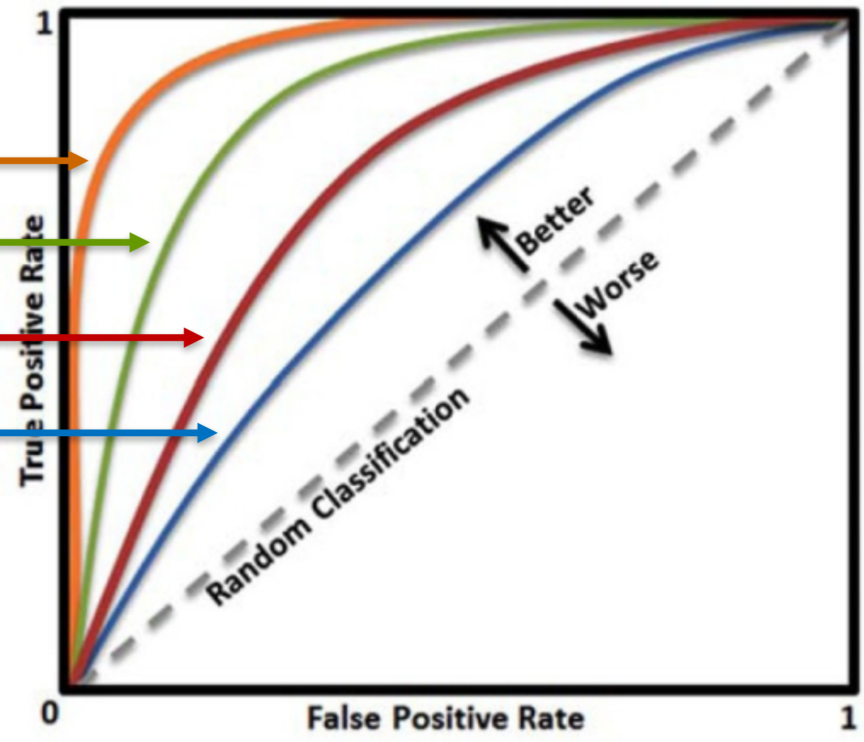
Area under the curve: 0.984

It means that in 98.4% of the time, a buyer will have a higher purchase probability than non-buyer.



ROC Values

- ROC Values
  - >0.9: excellent
  - 0.8-0.9 Good
  - 0.7-0.8 Fair
  - 0.6-0.7 Poor



Impact of Increasing Monetary Value by $1 on Purchase Probability

- Compute new utility of purchase

$$V_{new} = -30.29 + .111\text{Recency} + .594\text{Frequency} + .168(\text{Monetary}+1)$$

- Compute new probability of purchase

$$p_{new} = \frac{\exp(V_{new})}{\exp(V_{new}) + 1}$$

- Lift

$$\text{Lift} = \frac{p_{new} - p_{base}}{p_{base}}$$

Impact of Increasing Monetary Value by $1 on Purchase Probability

```
# calculate new logit probabilities (Monetary+1)
RFMdata_new <- RFMdata
RFMdata_new$Monetary <- RFMdata_new$Monetary + 1
RFMdata$New.Probability <- predict(model, RFMdata_new, type="response")
```

	Recency	Frequency	Monetary	Purchase	Base.Probability	New.Probability
1	120	7	41.66	0	0.0030728	0.0036319
2	90	9	46.71	0	0.0008332	0.0009852
3	120	6	103.99	1	0.9833225	0.9858611
4	270	17	37.13	1	0.9999999	0.9999999
5	60	5	88.92	0	0.0032378	0.0038267

Impact of Increasing the Monetary Value by $1 on Purchase Probability

- Avg. base purchase probability=0.45
- Avg. new purchase probability=0.45789
- Lift=(0.45789-0.45)/0.45=1.75%

Uses of Logistic Regression

- Rank customers from highest to lowest on a probability scale. Target those clients who:
  - Are at the top X% (“Customer Management/Allocation of Resources)
  - Who have probability above some cutoff (“Good Prospects”)
  - Who have slipped below some cutoff (About to “die” customers, Marketing Dashboard)
- Measure customers’ responsiveness to marketing actions
- Regression is not ok as you get estimates outside the range of [0,1] and wrong statistical tests