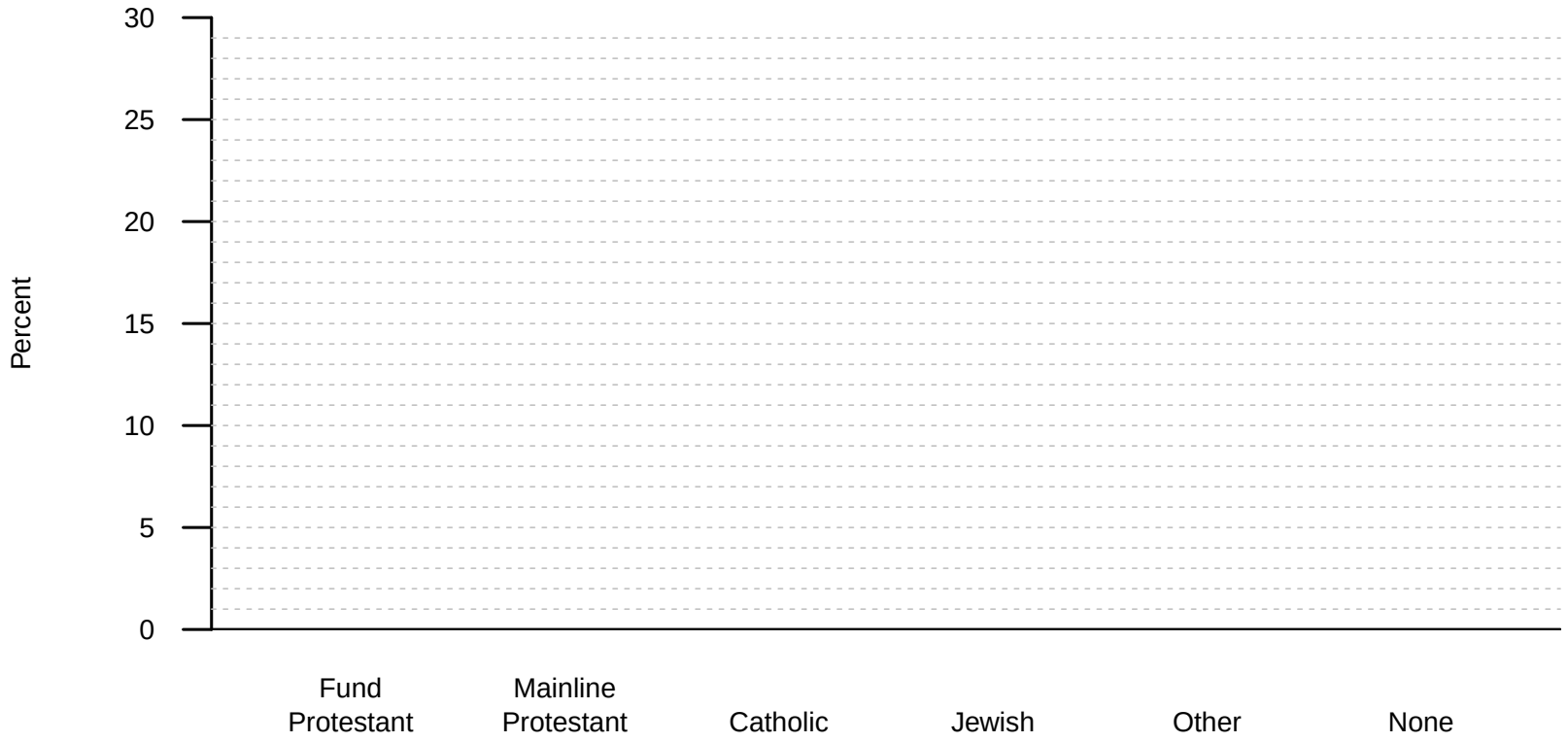
Handout 1: Draw a barplot of the distribution of religious affiliation

```
round(table(sex$relig)/sum(table(sex$relig)),2)*100
```

```
##
##     Fund Protestant Mainline Protestant            Catholic
##                  26                  25                  24
##              Jewish               Other                None
##                   2                   7                  16
```

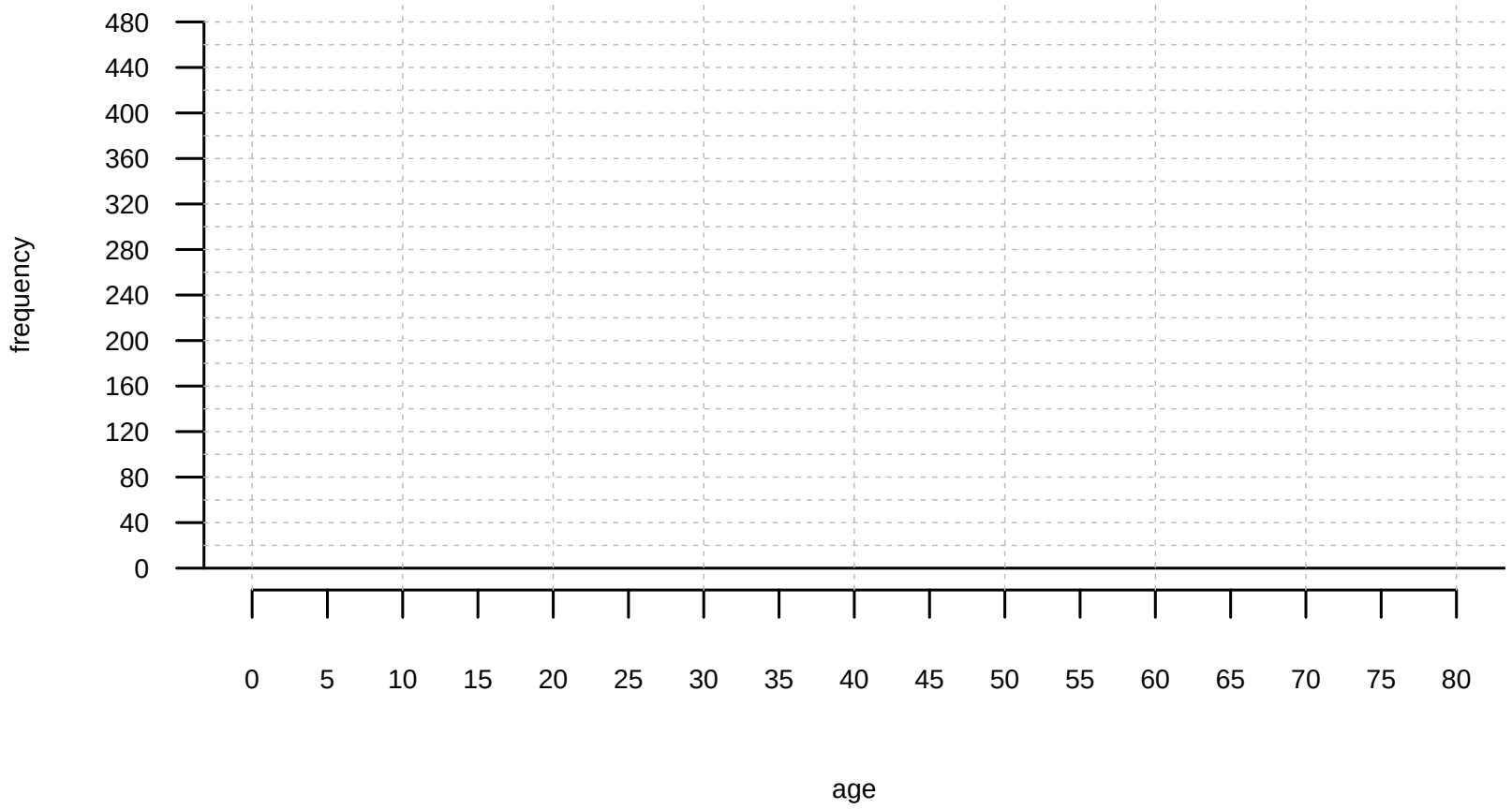


Name (Print and Sign): _____

#Handout 2: Draw a histogram of age on the Titanic

```
table(cut(titanic$age, breaks=seq(from=0, to=80, by=10)))
```

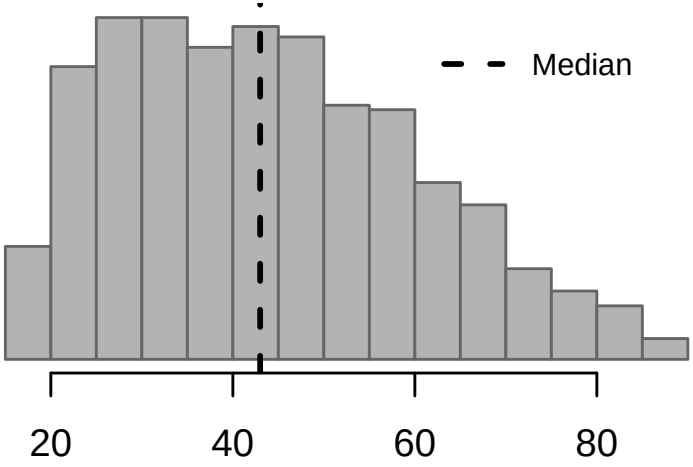
```
##
##  (0,10] (10,20] (20,30] (30,40] (40,50] (50,60] (60,70] (70,80]
##      97     208     470     266     161      71      30       6
```



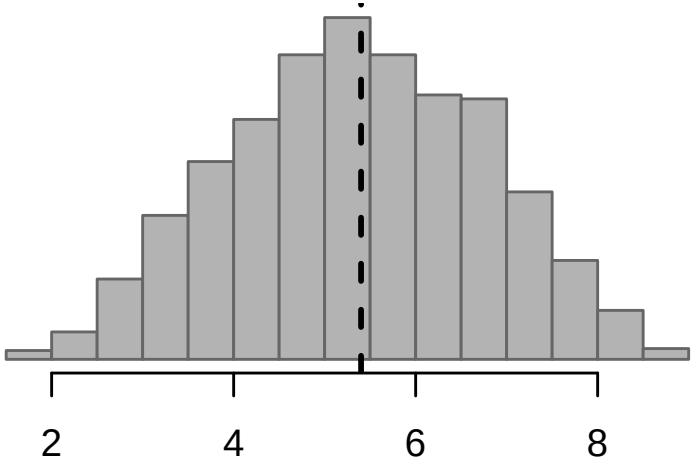
Name (Print and Sign): _____

Handout 3: Estimate and draw the balancing point of each distribution

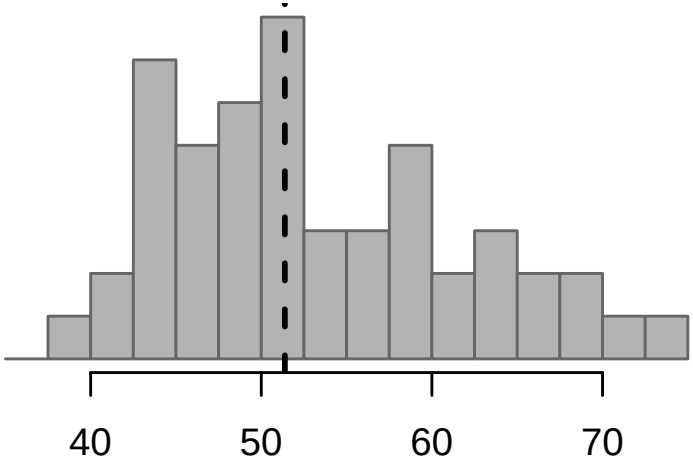
Age of respondents (sex survey)



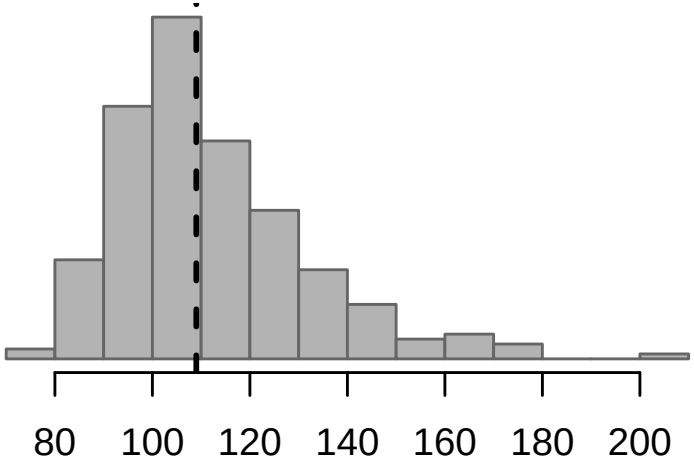
Movie tomato rating



Median income by state (crime data)



SciFi/Fantasy movie runtime

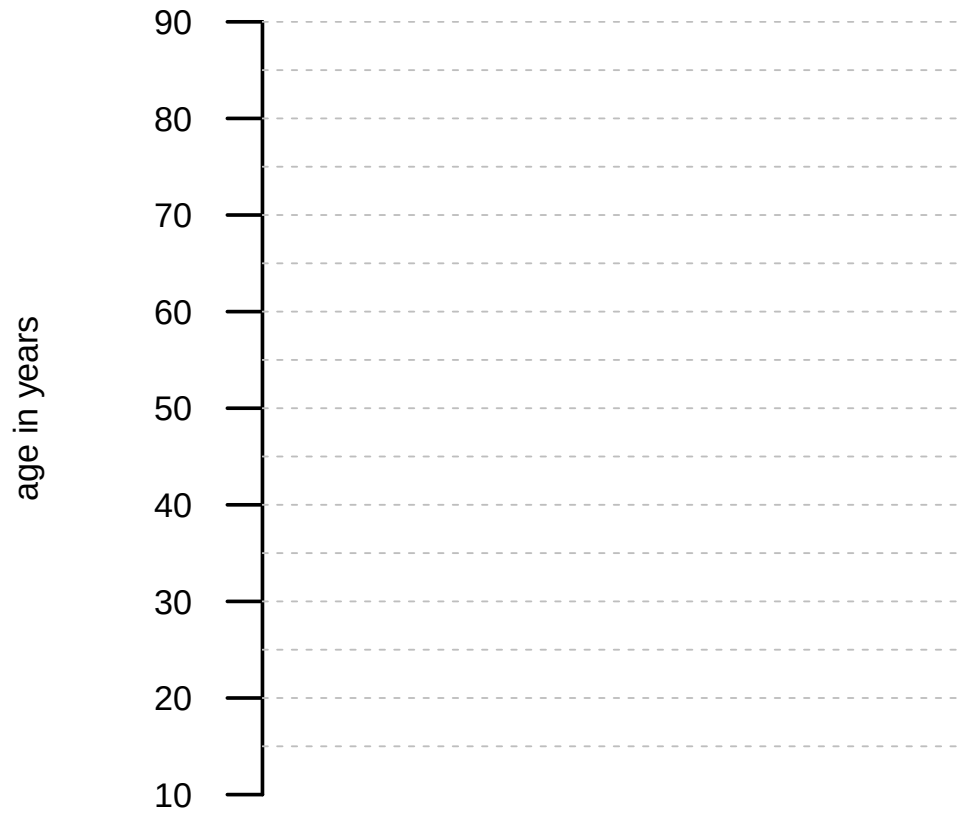


Name (Print and Sign): _____

Handout 4: Draw a boxplot of age distribution

```
quantile(sex$age)
```

```
##   0%  25%  50%  75% 100%
##   18   31   43   56   89
```



Name (Print and Sign): _____

Handout 5: Calculate variance and standard deviation of runtime for 2010 mystery movies

Movie	x	$x - \bar{x}$	$(x - \bar{x})^2$
All Good Things	101		
Edge of Darkness	117		
Wrecked	91		
Sum	309		

$$\bar{x} = 309/3 = 103$$

$$s_x = \sqrt{\sum_{i=1}^{n} (x - \bar{x})^2}/(3 - 1) =$$

Name (Print and Sign): _____

Handout 6: Calculate marginal distributions, conditional distribution of rating by genre, and odds ratio

Rating	**SciFi/Fantasy**	**Action**	**Total**
R or greater	66	106	
PG 13 or less	196	101	
Total			

Distribution of ratings for sciFi/fantasy movies:

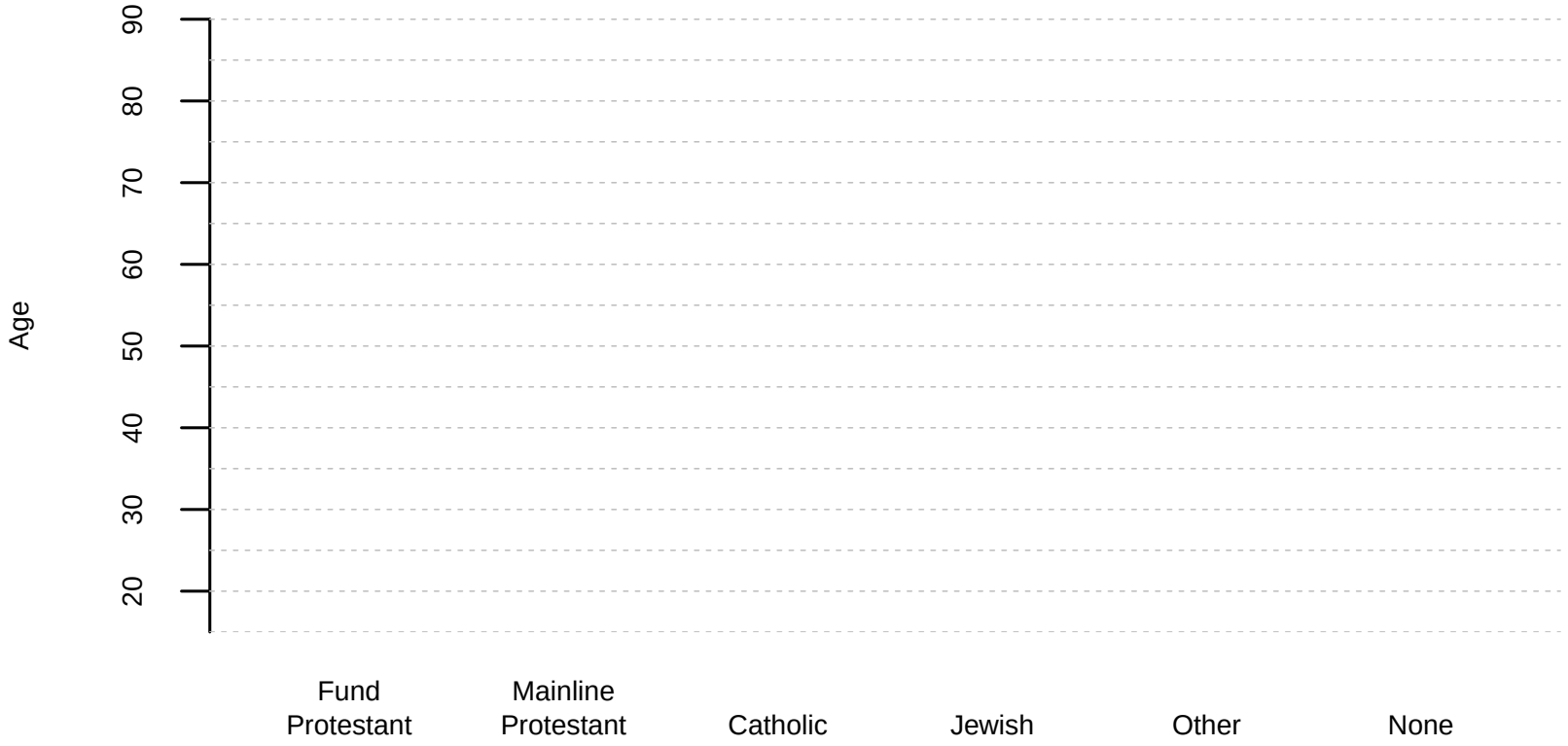
Distribution of ratings for action movies:

Odds ratio:

Name (Print and Sign): _____

Handout 7: Draw comparative boxplots of age by religious affiliation

	Fund P	Main P	Catholic	Jewish	Other	None
0%	18	18	18	21	18	18
25%	33	32	32	38	28	28
50%	44	46	43	53	37	37
75%	56	60	56	64	48	49
100%	89	89	88	89	77	85



Name (Print and Sign): _____

#Handout 8: Confidence interval for proportion supporting gay marriage

```
table(politics$gaymarriage)
```

```
##
## No legal recognition         Civil unions Support gay marriage
##                  776                  990                 2472
```

```
#sum up to get n
1543+1989+2382
```

```
## [1] 5914
```

```
#estimate p-hat
round(2382/5914,2)
```

```
## [1] 0.4
```

```
#t-stat
qt(0.975, 5914-1)
```

```
## [1] 1.960365
```

$$\texttt{standard error} = \sqrt{\hat{p} * (1 - \hat{p})/n} =$$

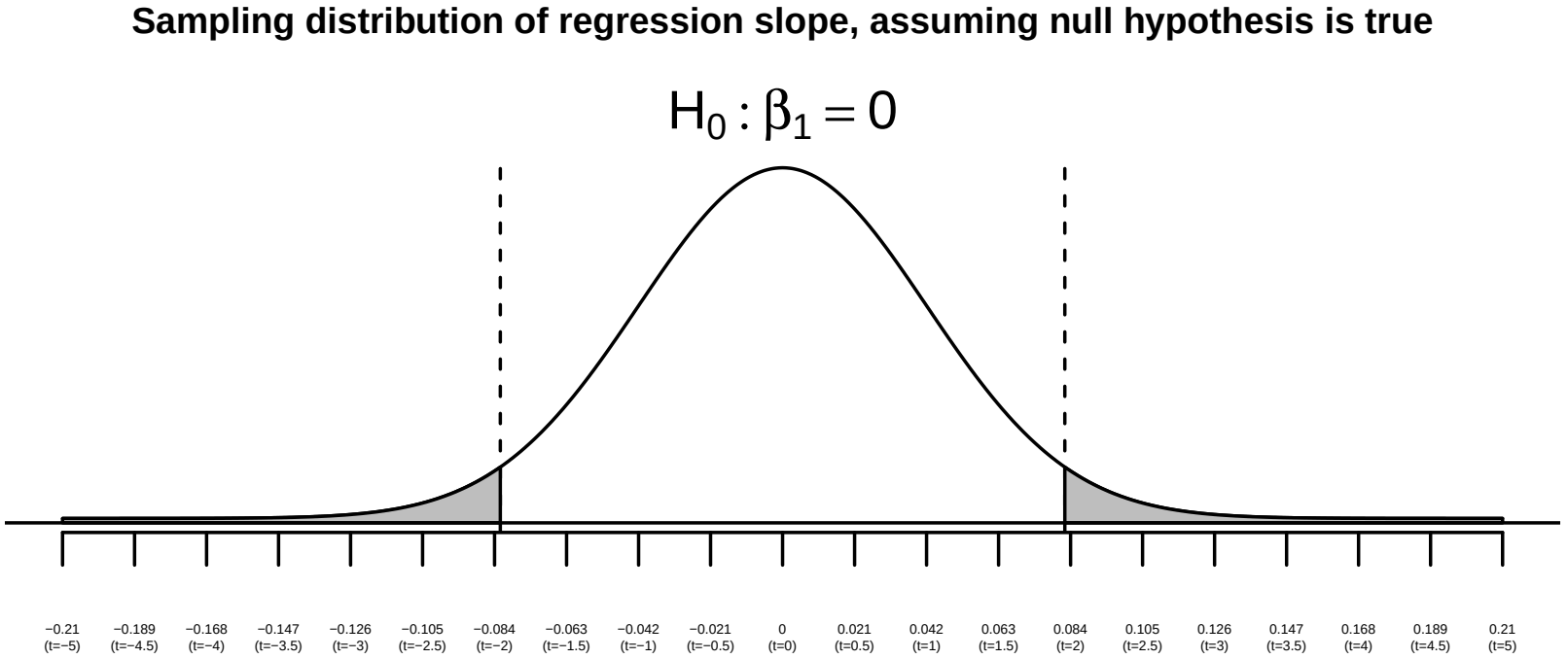
$$\texttt{confidence interval} = \hat{p} \pm t * (\texttt{standard error}) =$$

Name (Print and Sign): _____

Handout 9: Hypothesis test for a slope

```
round(summary(lm(income~age, data=politics))$coef[,1:2],3)
```

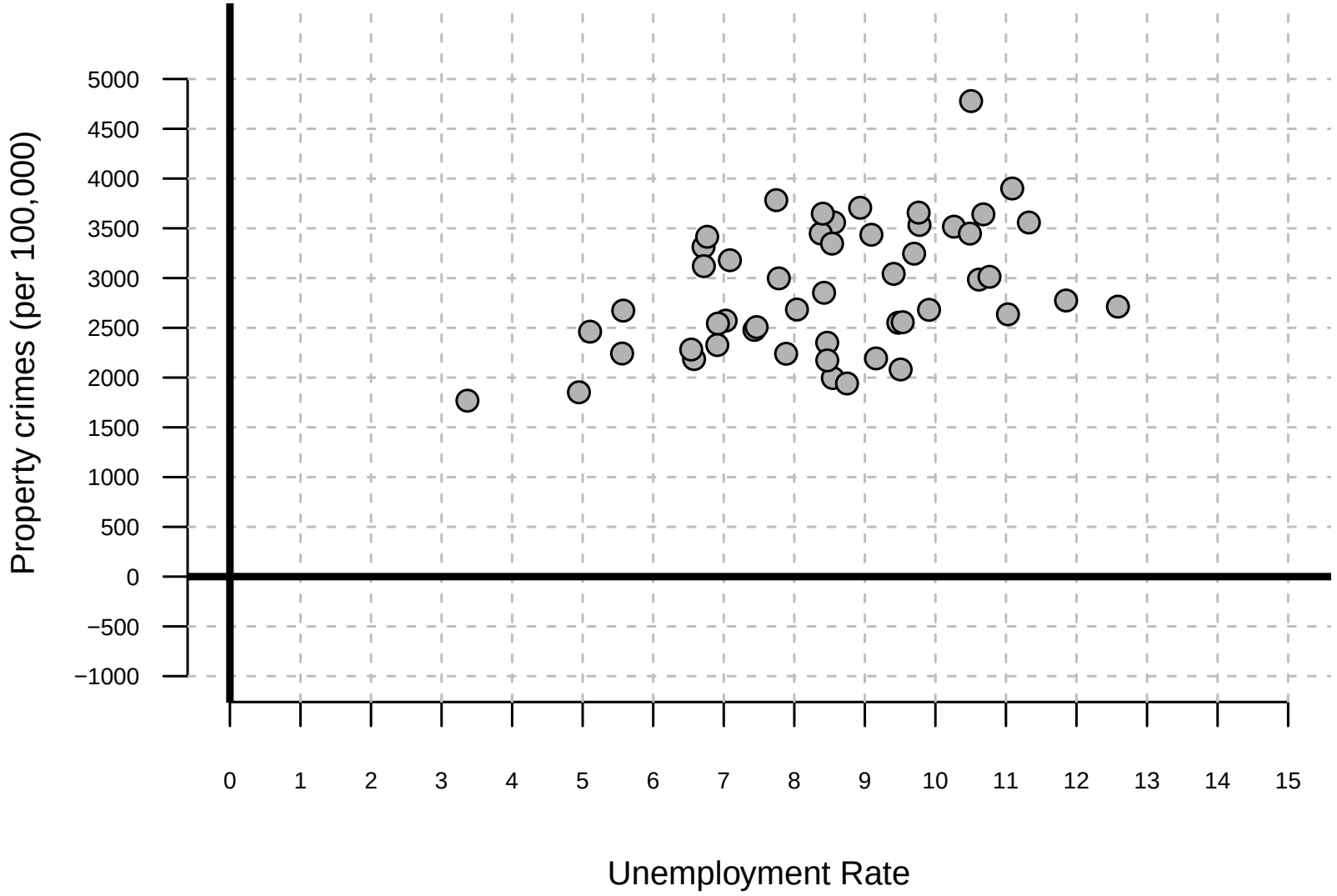
```
##             Estimate Std. Error
## (Intercept)   70.159      2.891
## age            0.051      0.055
```



Reject Fail to Reject

Name (Print and Sign): _____

Handout 10: Draw a straight line through the scatterplot and measure slope and intercept



Name (Print and Sign): _____

Handout 11: Interpret a slope and intercept

```
coef(lm(income~I(age-18), data=politics))
```

```
## (Intercept) I(age - 18)
## 71.06863240  0.05053182
```

Interpret the slope in a single sentence:

Interpret the intercept in a single sentence:

Name (Print and Sign): _____

Handout 12: Interpret a slope and intercept from a multivariate regression model

```
round(summary(lm(TomatoMeter~I(Year-2001)+I(Runtime-90)+I(BoxOffice-45), data=movies))$coef,3)
```

```
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         40.846      1.095  37.305    0.000
## I(Year - 2001)       0.318      0.135   2.358    0.018
## I(Runtime - 90)      0.324      0.032  10.227    0.000
## I(BoxOffice - 45)    0.059      0.008   7.349    0.000
```

Interpret the slope on year in a single sentence:

Interpret the slope on runtime in a single sentence:

Interpret the intercept in a single sentence:

Name (Print and Sign): _____

#Handout 13: Fill out the "slopes" for the regression models with different references

```
round(tapply(politics$income, politics$race, mean, na.rm=TRUE),0)
```

```
##                   White                   Black                  Latino
##                      78                      45                      57
## Asian/Pacific Islander        American Indian            Other/Mixed
##                     101                      38                      68
```

Variable	Estimate	Estimate
Intercept		
White	(reference)	
Black		(reference)
Hispanic		
Asian		
Am. Indian		

Name (Print and Sign): _____

Handout 13: Interpret the slope and intercept of regression model with categorical predictors

```
round(summary(lm(income~race+educ+I(age-25), data=politics))$coef,2)
```

```
##                              Estimate Std. Error t value Pr(>|t|)
## (Intercept)                     75.75       1.82   41.52     0.00
## raceBlack                      -25.21       3.07   -8.22     0.00
## raceLatino                     -10.05       2.97   -3.39     0.00
## raceAsian/Pacific Islander      13.14       4.84    2.72     0.01
## raceAmerican Indian            -23.79      11.06   -2.15     0.03
## raceOther/Mixed                 -8.15       4.45   -1.83     0.07
## educ.L                          61.26       2.75   22.27     0.00
## educ.Q                          13.33       2.43    5.49     0.00
## educ.C                           0.64       2.15    0.30     0.77
## educ^4                          -2.70       1.75   -1.54     0.12
## I(age - 25)                     -0.07       0.05   -1.34     0.18
```

Interpret the slope on Black in a single sentence:

Interpret the slope on BA degree in a single sentence:

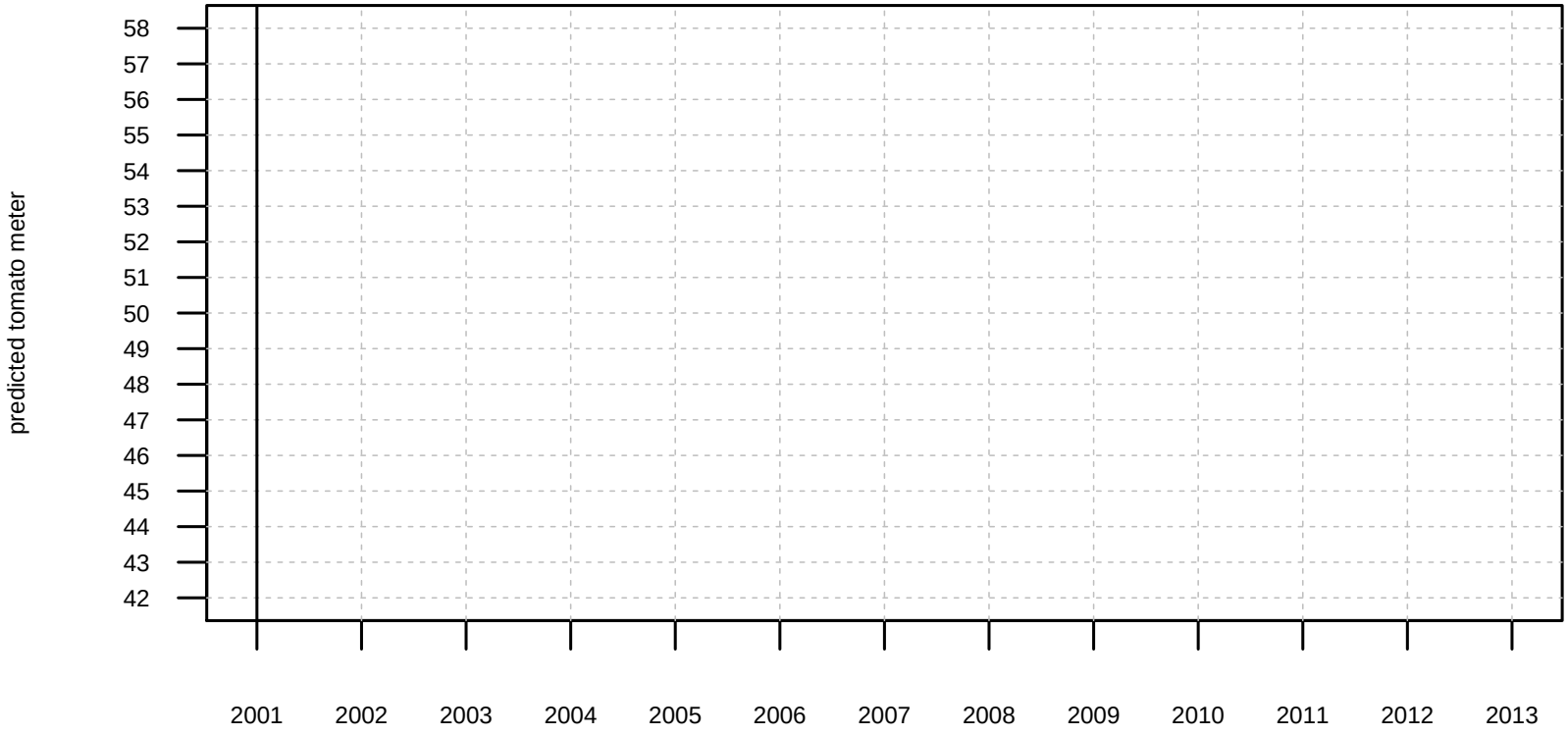
Interpret the intercept in a single sentence:

Name (Print and Sign): _____

Handout 15: Plot two lines showing movie ratings over time for popular and niche movies

```
movies$popular <- movies$Genre=="Action" | movies$Genre=="Animation" |
  movies$Genre=="Comedy" | movies$Genre=="Family" | movies$Genre=="SciFi/Fantasy"
round(summary(lm(movies$TomatoMeter~movies$popular*I(movies$Year-2001)))$coef,3)[,1:2]
```

```
##                                          Estimate Std. Error
## (Intercept)                                50.815      1.659
## movies$popularTRUE                         -8.246      2.099
## I(movies$Year - 2001)                       0.003      0.225
## movies$popularTRUE:I(movies$Year - 2001)    0.519      0.288
```



Name (Print and Sign): _____

Handout 16: Interpret the slope and intercept of regression model with interactions

```
round(summary(lm(BoxOffice~I(Runtime-90)*Rating, data=movies))$coef,2)
```

```
##                              Estimate Std. Error t value Pr(>|t|)
## (Intercept)                     73.99       7.69    9.62     0.00
## I(Runtime - 90)                  2.51       0.53    4.75     0.00
## RatingPG                       -14.64       8.53   -1.72     0.09
## RatingPG-13                    -56.26       8.12   -6.93     0.00
## RatingR                        -61.48       8.04   -7.65     0.00
## I(Runtime - 90):RatingPG        -1.65       0.57   -2.88     0.00
## I(Runtime - 90):RatingPG-13     -0.20       0.54   -0.37     0.71
## I(Runtime - 90):RatingR         -1.75       0.54   -3.25     0.00
```

Interpret the slope on Runtime in a single sentence:

Interpret the slope on RatingR in a single sentence:

Interpret the interaction term Runtime:RatingR in a single sentence:

Name (Print and Sign): _____