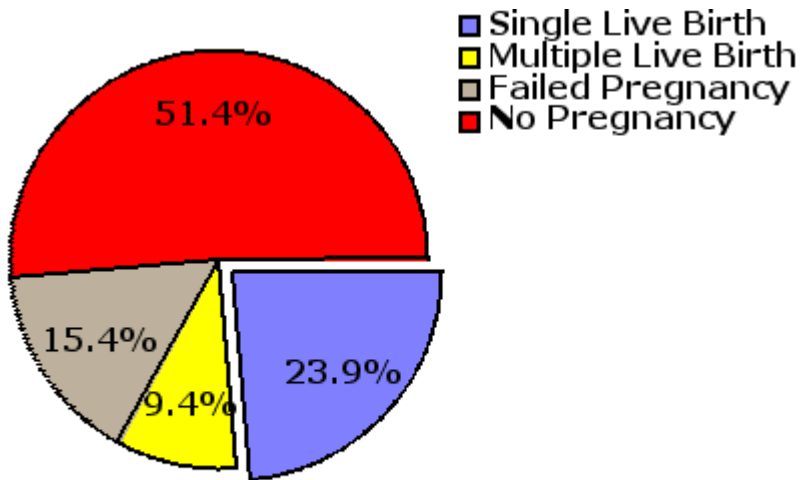
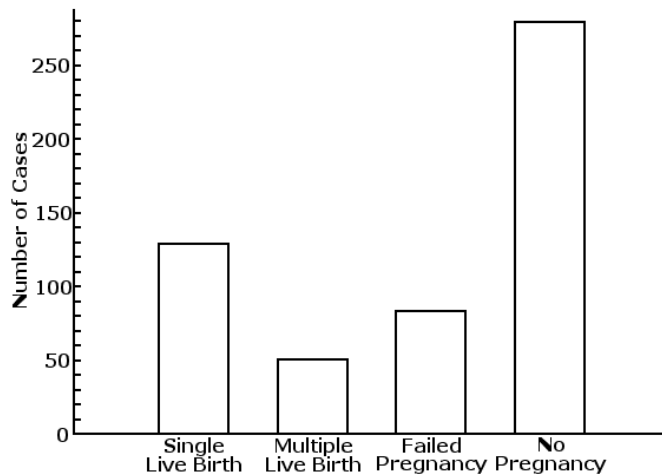


Example Answers for Assessment (Reproductive Medicine)

Question 1.1 : Pie Chart for Outcomes of IVF from Data in Worksheet A



Question 1.2 : Bar Chart for Outcomes of IVF from Data in Worksheet A



Question 1.3 : Sample Sizes and Precisions from Data in Worksheet A

Label	n	%	95%CI	ssiz ±1%	ssiz ±5%
Single_Live_Birth	130	23.9%	20.3% - 27.4%	6978	280
Multiple_Live_Birth	51	9.4%	6.9% - 11.8%	3259	131
Failed_Pregnancy	84	15.4%	12.4% - 18.4%	5009	201
No_Pregnancy	280	51.4%	47.2% - 55.8%	9597	384

Question 1.4 : Mean and SD of Endometrial Thickness from Worksheet B

n	56
mean	10.8
SD	2.3
df	55
SE	0.3
95% CI measurement	6.1 to 15.5
95% CI Error of mean	±0.6
95% CI mean	10.2 to 11.4

Question 1.5 : Percentile Analysis of Endometrial Thickness from Worksheet B

Percentile	Value
5 th Percentile	6.9
10 th Percentile	7.8
25 th Percentile	9.2
75 th Percentile	12.4
90 th Percentile	13.8
95 th Percentile	14.7

Question 1.6 : Sample Size for Endometrial Thickness from Results in Q1.4

±95% Error of mean	Sample Size
0.3	235
0.6	61
1.2	17

Question 1.7 : t and Percentile for Endometrial Thickness from Worksheet B

value	t	Percentile
10.3	-0.22	41
6.1	-2.02	2
13.2	1.03	85
10.9	0.04	52
14.0	1.37	91
9.6	-0.52	30
13.4	1.11	86

13.0	0.94	82
6.4	-1.89	3
10.3	-0.22	41
12.6	0.77	78
11.2	0.17	57
10.4	-0.18	43
9.6	-0.52	30
9.8	-0.43	33
10.9	0.04	52
10.9	0.04	52
5.5	-2.28	1
8.9	-0.82	21
10.2	-0.26	40
14.6	1.63	95
11.3	0.21	58
10.4	-0.18	43
14.4	1.54	94
10.2	-0.26	40
12.5	0.73	76
12.0	0.51	69
8.9	-0.82	21
9.4	-0.60	27
13.1	0.98	84
11.9	0.47	68
12.4	0.68	75
12.5	0.73	76
9.4	-0.60	27
10.9	0.04	52
11.0	0.08	53
11.2	0.17	57
10.3	-0.22	41
11.9	0.47	68
6.6	-1.81	4
13.8	1.28	90
8.6	-0.95	17
13.8	1.28	90
10.3	-0.22	41
11.1	0.12	55
11.9	0.47	68

13.7	1.24	89
10.1	-0.30	38
12.1	0.55	71
8.7	-0.90	18
16.7	2.53	99
6.8	-1.72	5
8.9	-0.82	21
11.8	0.43	66
6.7	-1.76	4
8.2	-1.12	13

Question 1.8 : Parametric Correlation (Pearson) Between Mercury Concentration and Sperm Count from Data in Worksheet C

n	52
Pearson's Correlation Coefficient (ρ)	-0.2744
Fisher's Z	-0.2816
Standard Error of Z	0.1429
95% CI ρ (One tail)	<-0.0459
95% CI ρ (Two tail)	-0.5092 to -0.0016

The 95% confidence interval (one tail, left) does not overlap the null value (0), so it is statistically significant. Mercury level and sperm counts are significantly negatively correlated, the higher the Mercury level, the lower the sperm count.

Question 1.9 : Sample Size for Parametric Correlation

ρ	Sample size for 1 tail model	Sample size for 2 tail model
-0.2744	80	101
-0.1372	327	414
-0.0915	737	935

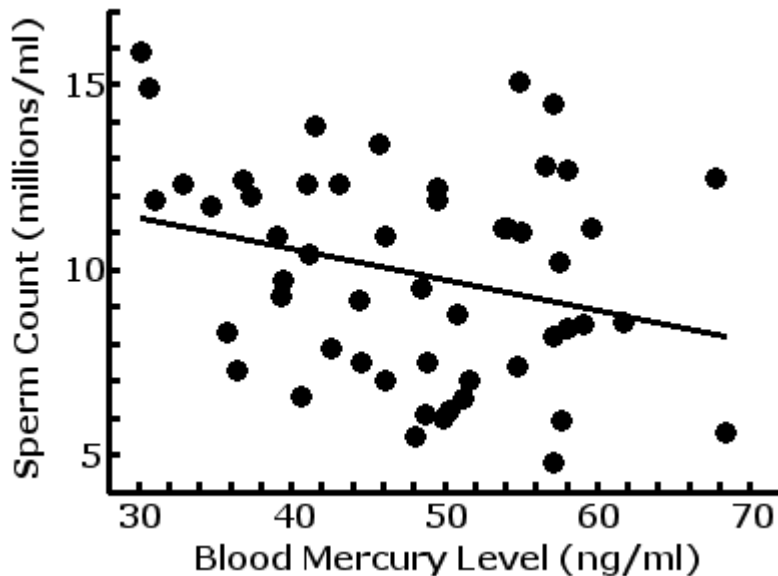
Question 1.10 : Regression Analysis Between Gestation and Birth Weight from Data in Worksheet C

Sperm count (millions / ml) = 13.8823 + -0.083(Mercury Concentration (ng / ml))

Mercury(ng/ml)	Sperm Count (millions / ml)
45	10.1

50	9.7
55	9.3
60	8.9
65	8.5

Question 1.11 : Scatter Plot Between Mercury Concentration and Sperm Count from Data in Worksheet C



Question 1.12 : Nonparametric Correlation between Sperm Motility (rows_ and percent of Malformed Sperms (columns) from Data in Worksheet D

	1:0%-10%	2:10%-30%	3:30%-70%	4:70%-100%
A: Linear Progressions	1	2	2	0
B: Non-linear progressions	1	9	8	1
C: Non-progressive Movements	4	3	6	1
D: No Movements	0	4	5	4

Spearman Correlation Coefficient: $n=51$ $r=0.2375$ $z=1.6794$ $p(1\text{ tail})=0.0465$. A significant negative correlation exists between sperm motility and percentage of sperms with abnormal morphology

Question 1.13 : Difference in Birth Weight of Babies from IVF Using Different Culture Media Using Data in Worksheet E

	n	Mean Nirth Weight (g)	Standard Deviation (g)
--	---	-----------------------	------------------------

Culture Medium A	23	3519	343
Culture Medium B	24	3471	322

Comparing Two Groups

Difference (meanA - meanB) = 48g Standard Error of Difference=97g

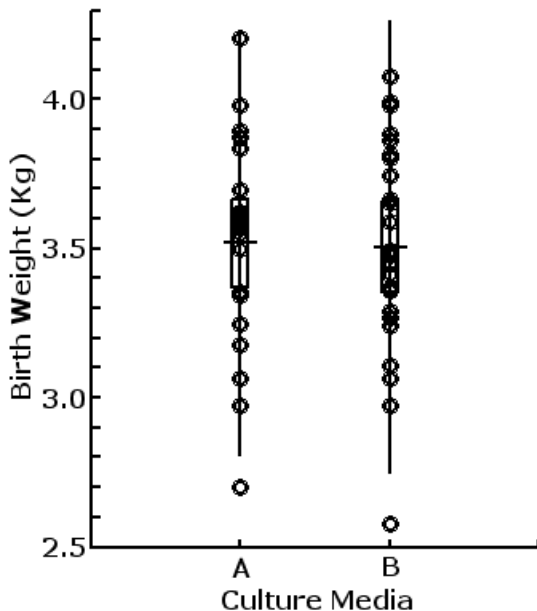
95% confidence interval of difference (2 tail) = -147g to +243g. Confidence interval overlaps the null (0) value, no significant difference detected.

The research hypothesis, that using different culture media may result in babies of different birth weight, is not supported by this set of results.

Question 1.14 : Sample Size comparing Two Parametric Measurements Based on Results of Q1.14

Difference to Detect	Expected Standard Deviation	Sample Size per group (1 tail)	Sample Size per group (2 tail)
48	350	659	836
48	400	860	1092
48	450	1088	1381

Question 1.15 : Difference in Birth Weight of Babies from IVF Using Different Culture Media Using Data in Worksheet E



The horizontal lines represents mean

The vertical lines represent 95% confidence interval of measurements

The boxes represents 95% confidence interval of the mean value

Question 1.16 : Comparison of Two Nonparametric Scales, Age of Mothers in 4 Groups between Success and Failure in Artificial Insemination, Using Data from Worksheet F

	1.Success with Live Birth	2.Failed with No Live Birth	Row Total
1:Age<30	6(23.1%)	2(8.3%)	8(16.0%)
2:Age=30-34	10(38.5%)	3(12.5%)	13(26.0%)
3: Age=35-39	9(34.6%)	16(66.7%)	25(50.0%)
4: Age=40+	1(3.8%)	3(12.5%)	4(8.0%)
Col Total	26(100.0%)	24(100.0%)	50(100.0%)

Mann-Whitney U Test : $U = -3.4256$ $p=0.0003$

A statistically significant difference exists. A negative U value means the averaged ranks in the first group (success) is lower than the second group (failure). In other words, the success group have younger women. These results support the hypothesis that success rate using artificial insemination decreases with increasing age of the patient

Question 1.17 : Risk Difference in Success with Live Birth Following IVF Between Using Hormonal Support Against Control, Using Data from Worksheet G

	1:Hormonal Support	2:Control	Row Total
+(Success with Live Birth)	14(56.0%)	11(44.0%)	25(50.0%)
-(Fail with no Live Birth)	11(44.0%)	14(56.0%)	25(50.0%)
Col Total	25(100.0%)	25(100.0%)	50(100.0%)

$Risk_{Hormone\ Support} = 0.56$ $Risk_{Control} = 0.44$

Risk Difference = 0.12 Standard Error (SE) = 0.1404

95% Confidence Interval of Difference (1 tail) = >-0.1109

95% Confidence Interval of Difference (2 tail) = -0.1552 to 0.3952

Number Needed to Treat (NNT) = 9

As we are interested only in whether hormonal support increases success, the one tail model is appropriate. The 95% confidence interval >-0.11 overlaps the null (0) value, so the risk difference is statistically not significant. The research hypothesis, that hormonal support following IVF improves success rate is not supported by this set of data and results.

Question 1.18 : Odds Ratio to Analyse a Number of Retrospective Studies Linking Smoking with Success in IVF (Live Birth), Using Data from Worksheet H

LB+ SM+	LB- SM+	LB+ SM-	LB- SM-	Odd of SM In LB+	Odd of SM In LB-	LOR	SE	OR	95% CI 1 tail	95% CI 2 tail	Support Hypothesis
176	227	25	6	7.04	37.8333	-1.6816	0.4656	0.1861	<0.4002	0.0747 to 0.4634	yes
84	85	5	8	16.8	10.625	0.4582	0.5905	1.5812	>0.5986	0.497 to 5.0303	no
165	172	10	8	16.5	21.5	-0.2647	0.4867	0.7674	<1.7089	0.2956 to 1.9921	no
162	142	13	13	12.4615	10.9231	0.1318	0.4087	1.1408	>0.5824	0.5121 to 2.5418	no
177	174	19	12	9.3158	14.5	-0.4424	0.3839	0.6425	<1.208	0.3028 to 1.3633	no
193	182	22	3	8.7727	60.6667	-1.9337	0.6241	0.1446	<0.4036	0.0426 to 0.4914	yes
226	219	30	4	7.5333	54.75	-1.9834	0.5407	0.1376	<0.3348	0.0477 to 0.397	yes
97	99	11	6	8.8182	16.5	-0.6265	0.5272	0.5344	<1.2721	0.1902 to 1.502	no

LB+SM+ = Success with live birth and smoked, LB-SM+ = Failure with no live birth and smoked, LB+SM- = Success with live birth and did not smoke, LB-SM- = Failure with no live birth and did not smoke.

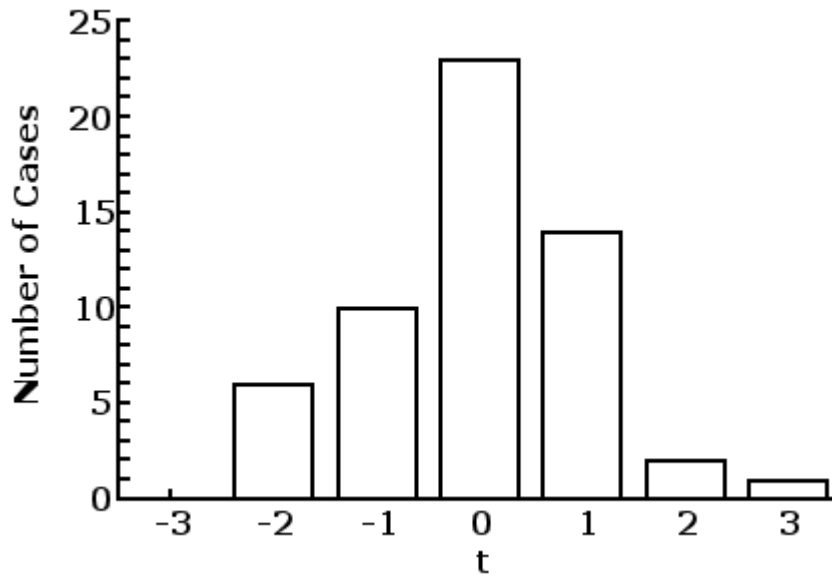
Odd of DM in LB+ = Odd of smoking in those that were success with live births. Odd of DM in LB- = Odd of smoking in those that were failures with no live births.

Log(OR) = Log(Odds Ratio), SE = Standard Error of Log(Odds Ratio)

OR = Odds Ratio

95% CI OR = 95% confidence interval of Odds Ratio = $\text{Exp}(\text{Log}(\text{OR}) \pm z(\text{SE}))$, where $z=1.65$ for one tail and 1.96 for two tail

Question 2.1.: Distribution Plot of Endometrial Thickness from Worksheet C



Question 2.2 : Meta-analysis of Log(OR) and its SE in results from Q1.18

Heterogeneity : Q Test : $Q=23.25$, $p=0.0015$, $I^2=69.9\%$. Heterogeneity Suspected. From I^2 , Heterogeneity can be considered Moderate

Publication Bias : Rank Correlation $z = -0.74$ $p = 0.229$. Publication Bias Not Suspected

Estimating Combined Summary Effect (Random Effect Model) :

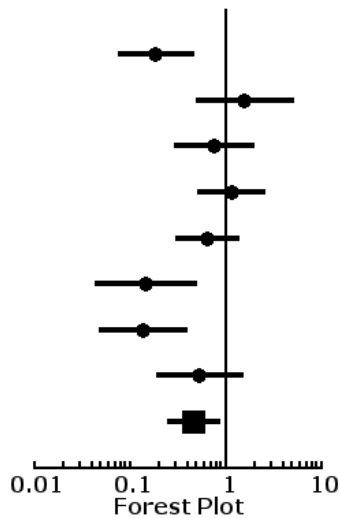
Study	Effect size Log(Odds Ratio)	Standard Error	95% CI of Log(Odds Ratio)	
1	-1.6816	0.4656	-2.5941	-0.7691
2	0.4582	0.5905	-0.6992	1.6155
3	-0.2647	0.4867	-1.2186	0.6892
4	0.1318	0.4087	-0.6693	0.9329
5	-0.4424	0.3839	-1.1948	0.31
6	-1.9337	0.6241	-3.1569	-0.7106
7	-1.9834	0.5407	-3.0432	-0.9237
8	-0.6265	0.5272	-1.6599	0.4069
REM	-0.766	0.3173	-1.3879	-0.1441

The combined Summary Effect size (Random Effect Model) is $\text{Log(OR)} = -0.766$, with 95% Confidence interval (two tail) of -1.388 to -0.144 . The 95% confidence interval does not overlaps the null (0) value, so it is statistically significant.

The conclusion is therefore that the results of the meta-analysis shows a significant negative relationship between smoking and success in IVF. This supports the research hypothesis that those with failed IVF had a greater odd of being smokers.

The 95% confidence interval of Log(OR) can translated to 95% confidence interval of Odds Ratio, so that the summary Odds ratio = $\exp(-0.766) = 0.465$, and the 95% confidence interval of summary oddes ratio = $\exp(-1.388)$ to $\exp(-0.144) = 0.250$ to 0.866

Question 2.3 : Forest Plot of Meta-analysis Results Obtained in Q2.2



The bottom most data point (square) is the summary effect size in Odds Ratio and its 95% confidence interval

Question 2.4 : Quality of Prediction Using Data in Worksheet I

	Success (LB+)	Failure (LB-)	Total
Previous Pregnancy (PP+)	40	47	87
Never Pregnant (PP-)	6	9	15
Total	46	56	102

Results

Parameters	Values
True Positive Rate	0.8696
False Positive Rate	0.8393
False Negative Rate	0.1304
True Negative Rate	0.1607
Likelihood Ratio for Test Positive	1.0361
Likelihood Ratio for Test Negative	0.8116

Question 2.5 : Post-test Probability of IVF Success with Live Birth, Based on Results in Q2.4

Pre-test Probability	Post-test Probability Previous Pregnancy+	Post-test Probability Previous Pregnancy-
25.0%	25.7%	21.3%
45.0%	45.9%	39.9%

Question 2.6 : Receiver Operator Characteristics for Duration of Infertility Predicting Success with Live Birth in IVF Treatment Using Data in Worksheet J

ROC=0.7124 SE=0.0527. 95% CI =0.609 to 0.81577

TPR=True Positive Rate=TP/(TP+FN)

TNR=True Negative Rate=TN/(TN+FP)=(1-FPR)

FNR=False Negative Rate=FN/(FN+TP)=(1-TPR)

YI=Youden Index=TPR - FPR

LR+=Likelihood Ratio Test Positive=TPR/FPR

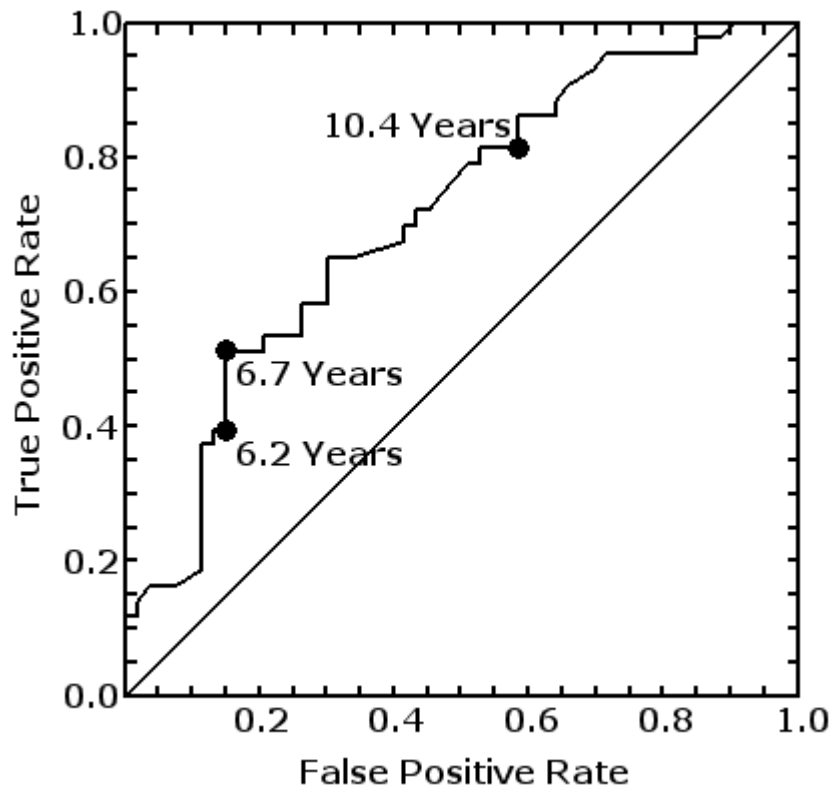
LR-=Likelihood Ratio Test negative=FNR/TNR

Value	FPR	TPR	TNR	YI	TPR/TNR	TNR/TPR	LR+	LR-
10.4	0.585	0.814	0.415	0.229	1.961	0.510	1.392	0.448
6.7	0.151	0.512	0.849	0.361	0.603	1.660	3.390	0.575
6.2	0.151	0.395	0.849	0.244	0.466	2.148	2.619	0.712

In the context of IVF

- 6.2 years of infertility, with the TNR/TPR closed to 2, marks where success is likely, but has no clinical utility so this figure need not be used
- 6.7 years of infertility, with a maximum Youden Index of 0.361, is a convenient cut off value for decision making. Patients with infertility for less than 6 years can be advised that the results of IVF is likely to be successful
- 10.4 years of infertility, with the TPR/TNR close to 2, marks the situation where success is unlikely. Patients with infertility for more than 10 years can be advised that success is unlikely and that they should consider abandoning IVF as treatment for their infertility

Question 2.7.: ROC Plot using Results Obtained in 2.6



Question 2.8.: Complex Post-test Probability Using Results from Q2.4 and Q2.6

Tests		Post-test Probability	
Previous Pregnancy	Years of Infertility	Pre-test Probability=25%	Pre-test Probability=45%
PP+	>6.7 yrs	16.6%	32.8%
PP+	<=6.7 yrs	54.0%	74.2%
PP-	>6.7 yrs	13.5%	27.6%
PP-	<=6.7 yrs	47.9%	69.2%

Question 2.9.: Covariance Analysis Comparing Duration of Infertility Between Success and Failure in IVF, corrected by the Covariate Age of the Patient, Using Data in Worksheet K

	Successful IVF (LB+)	Failed IVF (LB-)
--	----------------------	------------------

n	49	29
Mean Age (years)	36.4	35.7
SD Age	2.8	3.3
Mean Duration (years)	8.9	9.7
SD Duration	2.5	2.6
Slope b (years duration per year age)	0.5374	0.4723
Constant a (years duration when age=0)	-10.6	-7.2

Difference between unadjusted mean durations of infertility

Difference (Success-Failure) = -0.74 years duration

Standard Error = 0.5914

95% confidence interval (1 tail) = less than 0.25 years

95% confidence interval (two tail) = -1.9 to +0.4 years

Difference not statistically significant

Difference in slope

Difference (success-failure) = 0.0651 years duration per year age

Standard Error = 0.1603

95% confidence interval in difference between slopes (two tail) = -0.2544 to 0.3845

Difference in slopes not statistically significant

Adjusted difference in duration of infertility

Common slope (success and failure) = 0.5083 years duration per year age

Difference in duration of infertility, adjusted by common slope

Difference (success – failure) = -1.1 years

Standard Error = 0.48

95% CI of Difference (one tail) = less than -0.29 years

95% CI of Difference (two tail) = -2.05 years to -0.13 years

Summary and Conclusions

The duration of infertility are 8.9 years for successful IVF and 9.7 years for failed IVF. The difference is -0.74 years, not significantly different in both the one or two tail model

After adjustment by age, the difference is -1.1 years. This is statistically significant in both one and two tail model.

The conclusion is that patients with successful outcomes in IVF have shorter duration of infertility, once the duration is adjusted for age.

Question 2.10.: Covariance x/y Plot of Duration of Infertility, Success and Failure in IVF, and Age, Using Data in Worksheet K

