

1주차 3차시 : 회귀모형 개념

1. 회귀모형의 개념
2. 모집단 회귀함수
3. 표본 회귀함수

1.회귀모형의 개념

- 모집단 : 제주지역 물기업 60개
- Y(연간매출액, 단위 : 억 원)
- X(홍보비 지출액 , 단위:천 만원)

<표 2-1> 홍보비 지출액 및 연간매출액

X	80	100	120	140	160	180	200	220	240	260
연간 매출액	55	65	79	80	102	110	120	135	137	150
	60	70	84	93	107	115	136	137	145	152
	65	74	90	95	110	120	140	140	155	175
	70	80	94	103	116	130	144	152	165	178
	75	85	98	108	118	135	145	157	175	180
	-	88	-	113	125	140	-	160	189	185
	-	-	-	115	-	-	-	162	-	191
합계	325	462	445	707	678	750	685	1043	966	1211
E(Y X)	65	77	89	101	113	125	137	149	161	173

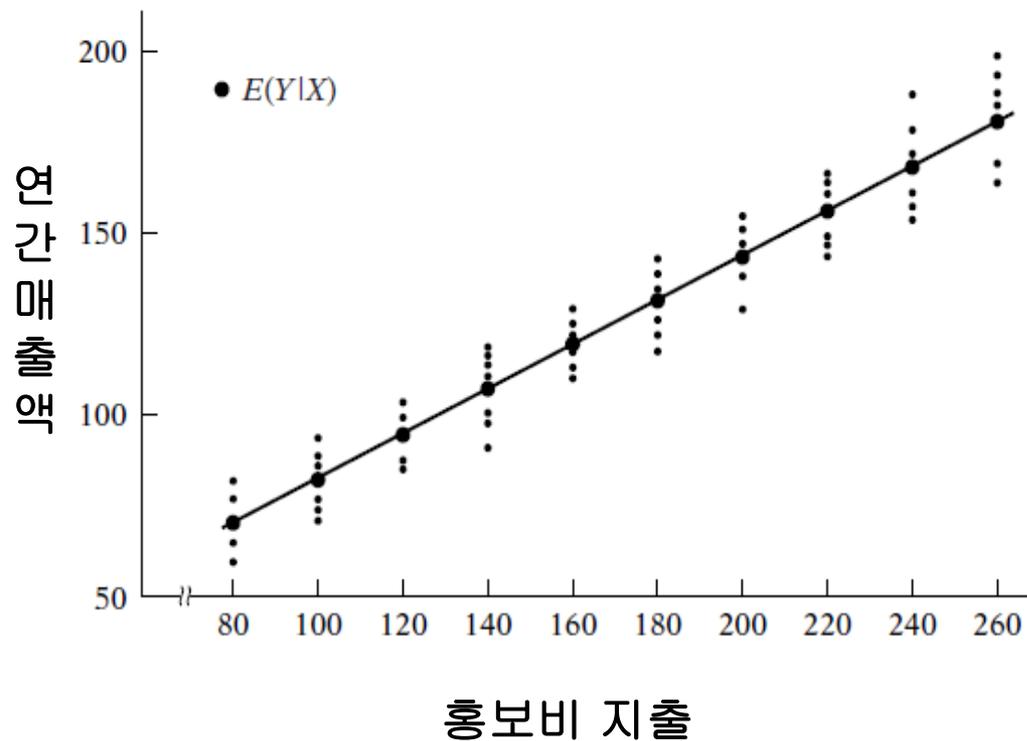
-X(홍보비 지출 수준)에 대한 Y(연간매출액)의 조건부 분포

<표 2-2> <표 2-1> 자료에 대한 조건부 확률

X Y	80	100	120	140	160	180	200	220	240	260
연 간 매 출 액	1/5	1/6	1/5	1/7	1/6	1/6	1/5	1/7	1/6	1/7
	1/5	1/6	1/5	1/7	1/6	1/6	1/5	1/7	1/6	1/7
	1/5	1/6	1/5	1/7	1/6	1/6	1/5	1/7	1/6	1/7
	1/5	1/6	1/5	1/7	1/6	1/6	1/5	1/7	1/6	1/7
	1/5	1/6	1/5	1/7	1/6	1/6	1/5	1/7	1/6	1/7
	-	1/6	-	1/7	1/6	1/6	-	1/7	1/6	1/7
	-	-	-	1/7	-	-	-	1/7	-	1/7
E(Y X)	65	77	89	101	113	125	137	149	161	173

-X(홍보비 지출 수준)에 대한 Y(연간매출액)의 조건부 분포

<그림 2-1> <표 2-1> 자료에 대한 조건부 분포



2.모집단 회귀함수

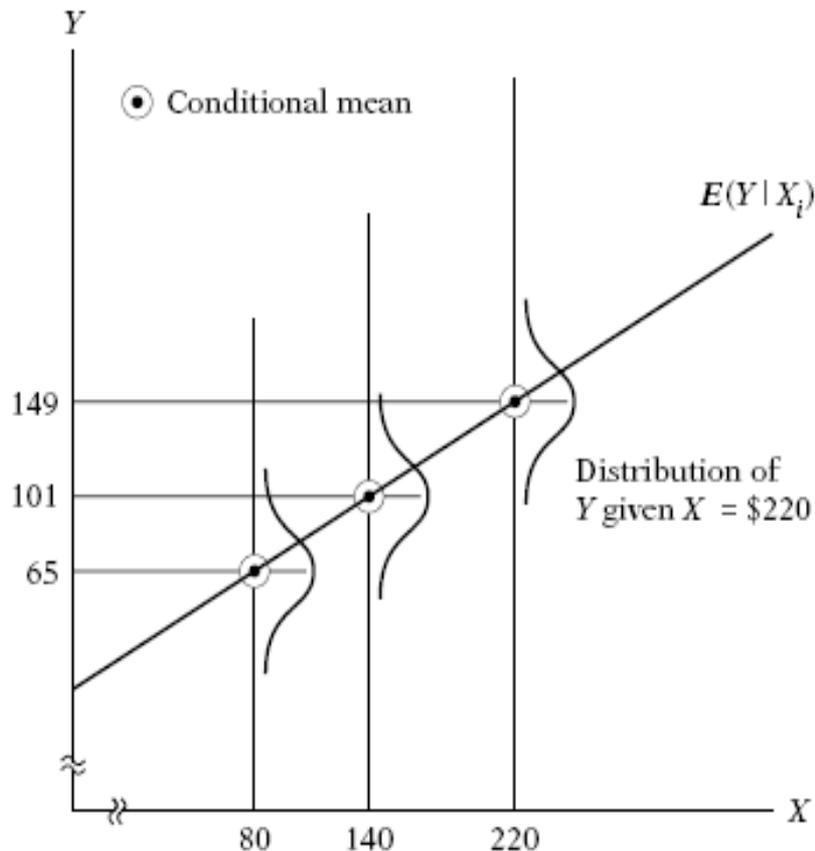
-모집단 회귀함수: $E(Y|X_i) = \beta_0 + \beta_1 X_i$

← 확정적(체계적) 요소

-모집단 회귀함수의 확률적 형태: $Y_i = E(Y|X_i) + u_i$ ← 확률적(비체계적) 요소

$$Y_i = \beta_0 + \beta_1 X_i + u_i$$

<그림 2-2> 모집단 회귀함수



3. 표본 회귀함수

<표 2-1>에서 추출한 확률표본 1

Y	X
70	80
65	100
90	120
95	140
110	160
115	180
120	200
140	220
155	240
150	260

<표 2-1>에서 추출한 확률표본 2

Y	X
55	80
88	100
90	120
80	140
118	160
120	180
145	200
135	220
145	240
175	260

- 표본회귀함수 : $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$

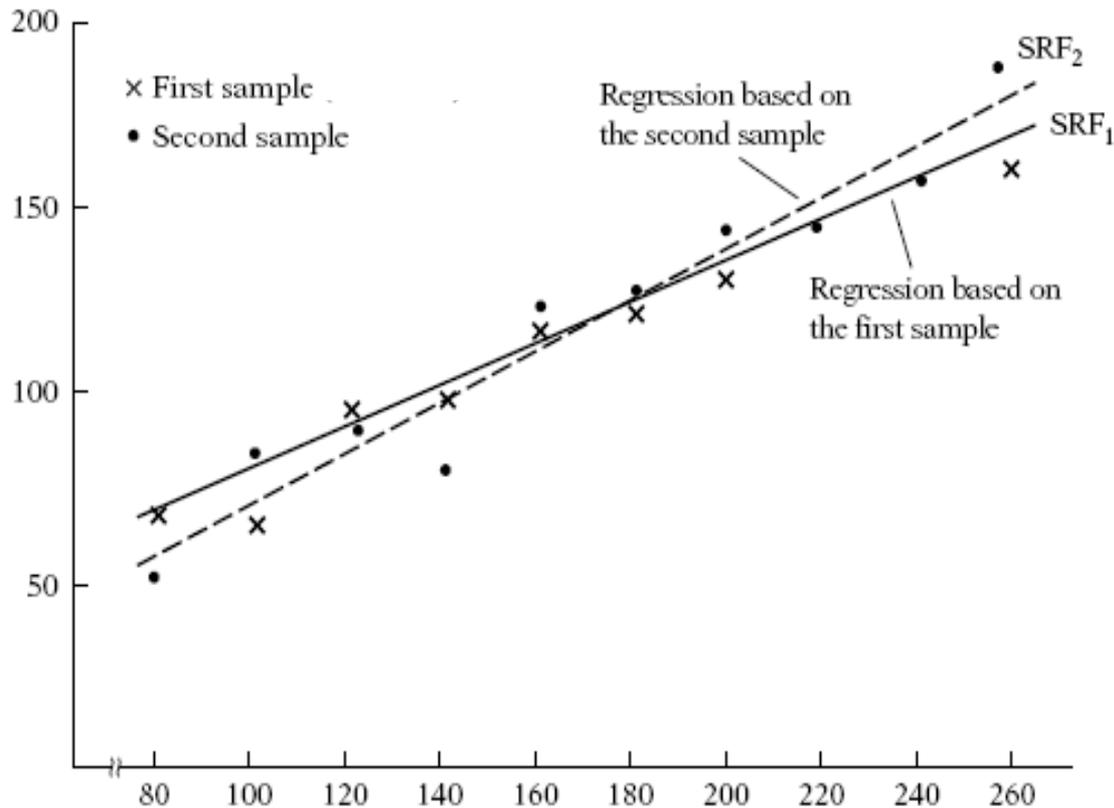
- 표본회귀함수의 확률적 형태 : $Y_i = \beta_0 + \beta_1 X_i + \hat{u}_i$

- \hat{Y}_i : $E(Y|X_i)$ 의 추정량

$\hat{\beta}_0$: β_0 의 추정량

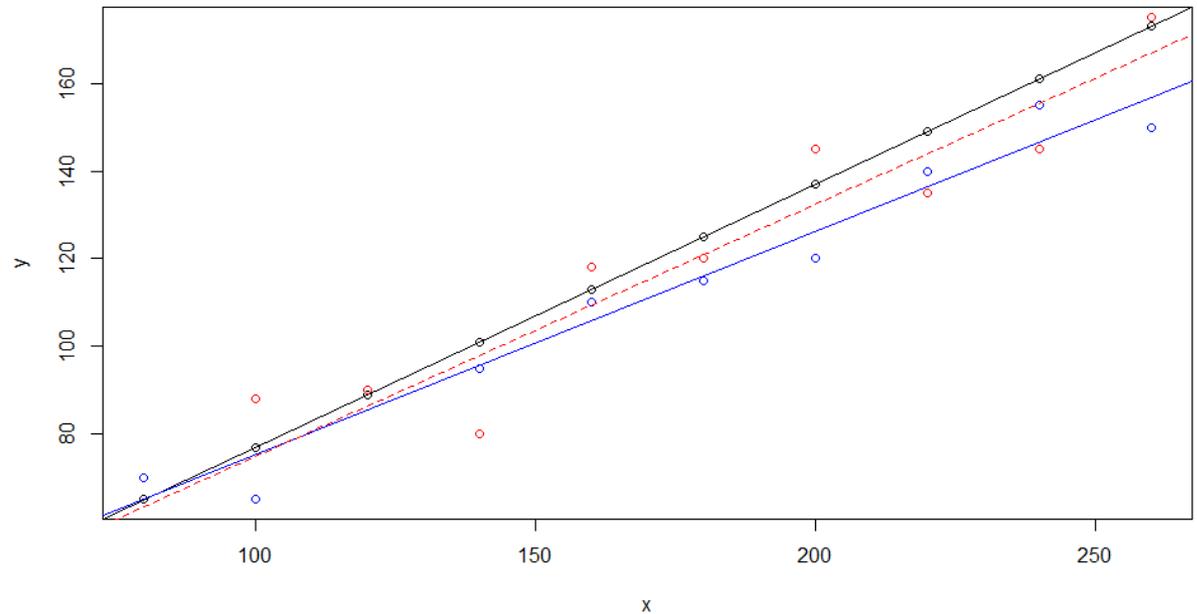
$\hat{\beta}_1$: β_1 의 추정량

<그림 2-3> 두 확률표본에 근거한 표본회귀함수(회귀선)

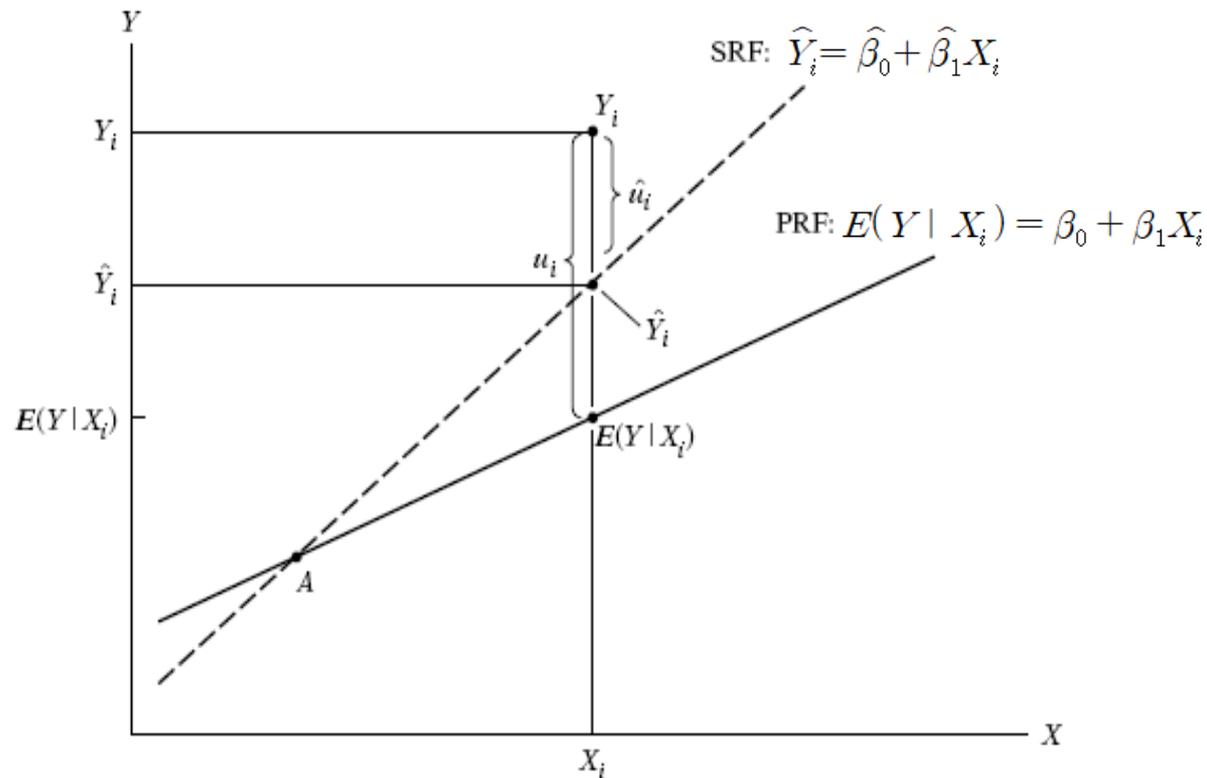


<그림 2-3>을 R로 그려보면 다음과 같음

```
x<-c(80,100,120,140,160,180,200,220,240,260)
y<-c(65,77,89,101,113,125,137,149,161,173)
y1<-c(70,65,90,95,110,115,120,140,155,150)
y2<-c(55,88,90,80,118,120,145,135,145,175)
plot(x,y,cex=1.0, col="black")
abline(lm(y~x),col="black")
points(x,y1,cex=1.0, col="blue")
points(x,y2,cex=1.0, col="red")
abline(lm(y1~x),col="blue")
abline(lm(y2~x),col="red",lty=2)
```



<그림 2-4> 모집단회귀함수와 표본회귀함수



결론 : 회귀분석은 표본회귀함수를 기초로 모집단회귀함수를 추정

$$\text{표본회귀함수 : } Y_i = \hat{\beta}_0 + \hat{\beta}_1 X_i + \hat{u}_i$$

↓ (추정)

$$\text{모집단회귀함수 : } Y_i = \beta_0 + \beta_1 X_i + u_i$$