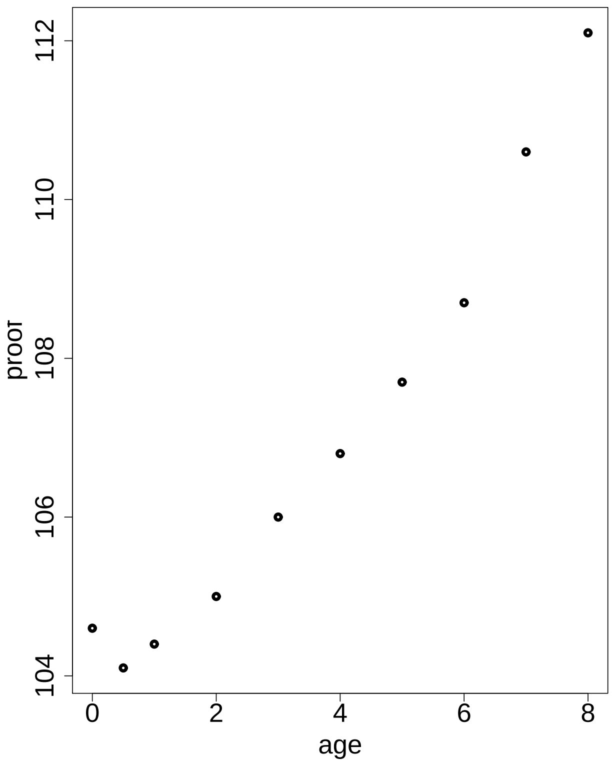
4. Simultaneous Inferences

Data: X_i = Age of Whiskey, Y_i = Proof

X	Y
0	104.6
0.5	104.1
1	104.4
2	105
3	106
4	106.8
5	107.7
6	108.7
7	110.6
8	112.1



Assumption of a SLR results in

$$\widehat{\mathsf{E}}(Y) = 103.5 + 0.955X$$

$r^2 = 0.9487$, $\sqrt{\mathsf{MSE}} = 0.6617$, $\bar{X} = 3.65$, $S_{XX} = 71.025$

We plan on selling 2 types of Whiskey:
- 2 years old
- 5 years old

Government requires CI's for the proof that hold **jointly** at 95%.
- Proof after 2 years: $\mathsf{E}(Y_2) = \beta_0 + \beta_1(2)$
- Proof after 5 years: $\mathsf{E}(Y_5) = \beta_0 + \beta_1(5)$

- 95% CI for $\mathsf{E}(Y_2)$ is

$$b_0 + b_1(2) \pm t(0.975; 8)\sqrt{\mathsf{MSE}\left(\frac{1}{10} + \frac{(2-\bar{X})^2}{S_{XX}}\right)} = (104.86, 105.99)$$

• 95% CI for $\mathsf{E}(Y_5)$ is

$$b_0 + b_1(5) \pm t(0.975; 8)\sqrt{\mathsf{MSE}\left(\frac{1}{10} + \frac{(5-\bar{X})^2}{S_{XX}}\right)} = (107.75, 108.83)$$

We say that a CI **covers**, if it contains the true parameter.

Define A_2 to be the event that the CI for $\mathsf{E}(Y_2)$ covers. Define A_5 to be the event that the CI for $\mathsf{E}(Y_5)$ covers.

Then

$$P(A_2) = 0.95, \quad P(A_5) = 0.95, \quad P(A_2 \cap A_5) \stackrel{?}{=} 0.95$$

NO! Let's do some experiments:

exp't #	A_2	A_5	$A_2 \cap A_5$
1	yes	yes	yes
2	yes	no	no
3	yes	yes	yes
4	no	yes	no
5	yes	yes	yes
6	yes	yes	yes
⋮	⋮	⋮	⋮

There will be 5% no's in the A_2 column and 5% no's in the A_5 column. But there are more than 5% (but less than 10%) no's in the $A_2 \cap A_5$ column.

$$
\begin{aligned}
P(A_2 \cap A_5) &= P(A_2) + P(A_5) - P(A_2 \cup A_5) \\
&= 1 - P(\overline{A_2}) + 1 - P(\overline{A_5}) - P(A_2 \cup A_5) \\
&= 1 - P(\overline{A_2}) - P(\overline{A_5}) + P(\overline{A_2 \cup A_5}) \\
&\geq 1 - \left(P(\overline{A_2}) + P(\overline{A_5})\right)
\end{aligned}
$$

$\overline{A_2}$ is the event that the CI for $\mathsf{E}(Y_2)$ does **not cover**, with $P(\overline{A_2}) = \alpha$.

Conclusion: To get $P(A_2 \cap A_5) \geq 0.95$ we need $P(\overline{A_2}) + P(\overline{A_5}) \leq 0.05$.

General Bonferroni Inequality: for g such events A_i

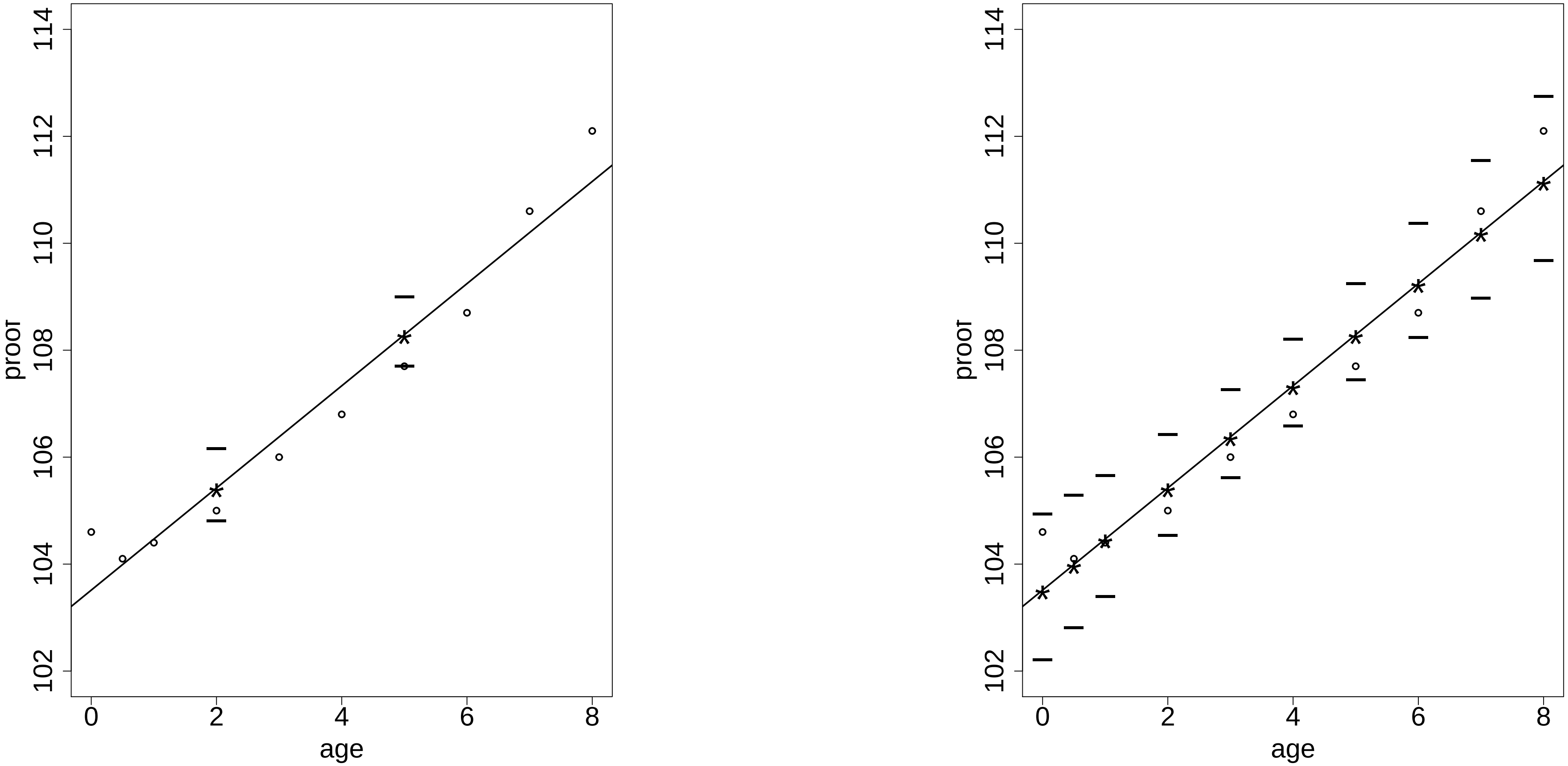
$$P(A_1 \cap A_2 \cap \ldots \cap A_g) \geq 1 - \sum_{i=1}^{g} P(\overline{A_i})$$

Thus, if $P(\overline{A_i}) = \alpha/g$ we get

$$P(A_1 \cap A_2 \cap \ldots \cap A_g) \geq 1 - \sum_{i=1}^{g} \frac{\alpha}{g} = 1 - \alpha$$

$$\begin{aligned}
\hat{Y}_1 \quad &\pm \quad t\left(1 - \frac{\alpha}{2g}; n - 2\right)\sqrt{\mathsf{MSE}\left(\frac{1}{n} + \frac{(X_1 - \bar{X})^2}{S_{XX}}\right)} \\
\hat{Y}_2 \quad &\pm \quad t\left(1 - \frac{\alpha}{2g}; n - 2\right)\sqrt{\mathsf{MSE}\left(\frac{1}{n} + \frac{(X_2 - \bar{X})^2}{S_{XX}}\right)} \\
&\vdots \\
\hat{Y}_g \quad &\pm \quad t\left(1 - \frac{\alpha}{2g}; n - 2\right)\sqrt{\mathsf{MSE}\left(\frac{1}{n} + \frac{(X_g - \bar{X})^2}{S_{XX}}\right)}
\end{aligned}$$

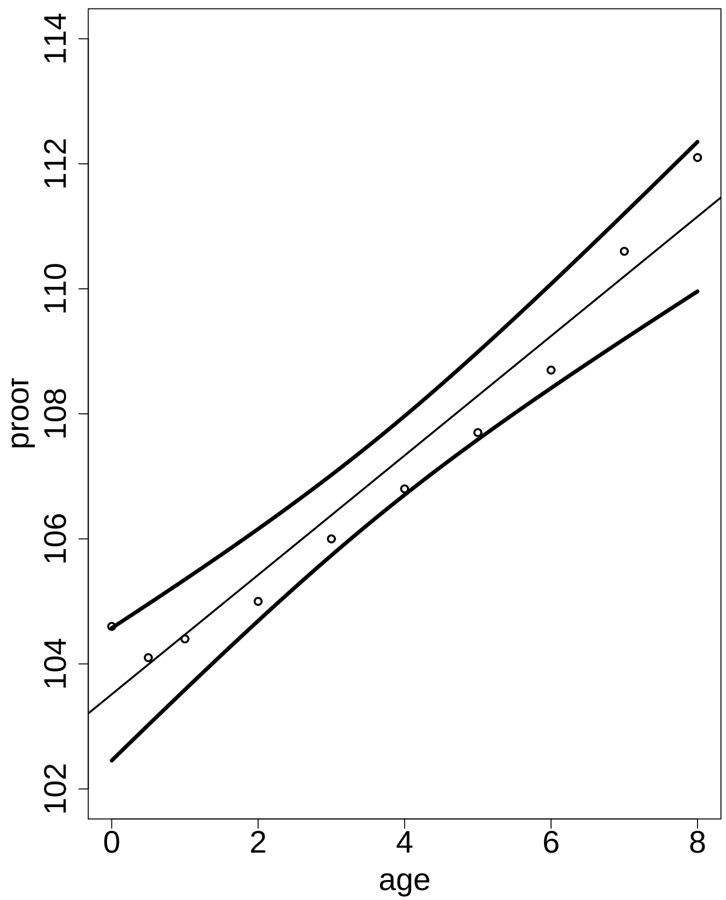
These CI's hold **jointly** at $(1-\alpha)100\%$.



Working-Hotelling Procedure: give confidence band for the entire (true) regression line, $\beta_0 + \beta_1 X$, within the range of the data.

This band is constructed pointwise at X_a

$$\hat{Y}_a \pm \sqrt{2F(1-\alpha; 2, n-2)}\sqrt{\mathsf{MSE}\left(\frac{1}{n} + \frac{(X_a - \bar{X})^2}{S_{XX}}\right)}$$



Thinnest at $\bar{X}$.
We are $(1-\alpha)100\%$ confident that the true regression function is within this band.
Since the whole band has $(1-\alpha)100\%$ confidence, you can pick as many CI's for E(Y) at as many different X's as you want and the **joint** confidence is at least $(1-\alpha)100\%$.

Comparing both CI's: It is simple to compare which one gives smaller intervals, just compare

$$t\left(1-\frac{\alpha}{2g};n-2\right) \quad \text{to} \quad \sqrt{2F(1-\alpha;2,n-2)}$$

Example: We want 2 CI's for the proof of Whiskey when age is 2 and 5 years to hold jointly at 95%.

Bonferroni: use 2 CI's each at level $\alpha/g = \alpha/2 = 0.025$

$$t(1-0.0125;8) = 2.751.$$

Working-Hotelling:

$$\sqrt{2F(0.95;2,8)} = 2.986.$$

	Individual CI's	Bonferroni joint	W-Hotelling
2	(104.86,105.99)	(104.75,106.10)	(104.69,106.16)
5	(107.75,108.83)	(107.64,108.94)	(107.59,108.99)