

Errata för ”Regression med R och Python - Introduktion och fördjupning”

senast uppdaterad 9 juni 2025

- s. 21, s.23: I figur 2.2 och 2.3 ska det stå r_{XY} i stället för r_{xy} .
- s. 41: I ekvation 3.13 ska det stå $P(\{Y = k\} \cap \{X = h\})$ i stället för $P(\{Y = k\}) \cap P(\{X = h\})$.
- s. 50: Näst sista stycket saknas hänvisning till figur 3.2. Det ska stå ”Resultatet blir dock annorlunda om Y följer en χ^2 -fördelning med en frihetsgrad. I figur 3.2 ser vi att...”
- s. 96: I övningsuppgift 5.3 ska det stå `apartment_price_data` i stället för `apartment_sales_data`.
- s. 112: I rutan har det blivit ett betydande fel där de två första punkterna ska vara tvärtom. Korrekt skrivning är

- Skattningen av koefficienten framför X_j kommer att **minska** när X_{K+1} läggs till i modellen, om riktningen på sambandet mellan X_j och X_{K+1} har samma tecken som riktningen på sambandet mellan X_{K+1} och Y .
 - Skattningen av koefficienten framför X_j kommer att **öka** när X_{K+1} läggs till i modellen, om riktningen på sambandet mellan X_j och X_{K+1} har annat tecken än riktningen på sambandet mellan X_{K+1} och Y .

- s. 113: I ekvation 6.31 och texten precis ovan ska det stå β_K och inte β_k .
- s. 142: I stycket ovanför ekvation 7.20 ska det stå $\hat{E}(\ln Y|X) = \hat{\beta}_0 + \hat{\beta}_1 X$ i stället för $\hat{E}(\ln Y|X) = \hat{\beta}_0 + \hat{\beta}_1$.
- s. 145: I ekvation 7.27 ska det stå Δx i stället för ΔX .
- s. 146: I ekvation 7.34 ska det stå D i stället för d .
- s. 147–156: Exemplet i avsnitt 7.5 gällande vindriktning och luftförorening är inspirerat av ett exempel i en artikel av Deryugina m.fl., 2019 som studerar samband mellan luftförorening och hälsa där luftförorening instrumenteras med vindriktning. Den artikeln använder data från en annan tidsperiod än det data som används i boken.
- s. 157: I övningsuppgift 7.1d ska det stå `xvals` i stället för `vals` i raden där `yhat` skapas. Dessutom måste observationer med saknade värden på byggnadsår tas bort för att koden

ska gå att använda. Detta kan exempelvis uppnås genom att – innan koden i uppgiften körs – skriva

```
df <- df[!is.na(df$build_year), ]
```

- s. 158: I övningsuppgift 7.2e ska det stå $\hat{E}(Y|X = 1981) - \hat{E}(Y|X = 1980)$ i stället för $\hat{E}(Y|X = 1991) - \hat{E}(Y|X = 1990)$.
- s. 159: I uppgift 7.5 ska det stå "intresserade" i stället för "intresserad".
- s. 161: Högst upp ska det stå "lägenheter" i stället för "lägeneter".
- s. 161: I övningsuppgift 7.8b beror det på vilken referenskategori för antal rum som används huruvida någon koefficient blir statistiskt signifikant.
- s. 188: Sista stycket 1:a raden: $E(Y = 1|X)$ ska i stället vara $E(Y|X)$.
- s. 196: I ekvation 9.23 gäller den betingade sannolikheten betingat på samtliga X -variabler och ekvationen ska därför skrivas som

$$P(Y = 1|X_1, \dots, X_K) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \dots + \beta_K X_K)}}.$$

- s. 196: I ekvation 9.24 ska oddskvoten vara $e^{\beta_1 \Delta x}$ och inte bara e^{β_1} .
- s. 196: I ekvation 9.25 har det råka komma med en extra etta. Korrekt uttryck är:

$$\text{Riskkvot, RK} : \frac{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_K X_K)}}{1 + e^{-(\beta_0 + \beta_1 (X_1 + \Delta x_1) + \beta_2 X_2 + \dots + \beta_K X_K)}}.$$

- s. 197: I ekvation 9.27 ska det stå Z i stället för z .
- s. 197: I ekvation 9.28 ska det stå $\hat{V}$ i stället för $\hat{V}$.
- s. 201: Högst upp på sidan ska det stå $y_i - P(Y = 1|X = x_i)$ i stället för $y_i - P(y_i = 1|x_i)$.
- s. 201: Längst ner på sidan ska det stå

```
logit_model <- glm(high_income ~ share_tertiary_school,
```

i stället för

```
logit_model <- glm(df$high_income ~ share_tertiary_school,
```

- s. 218: Övningsuppgift 9.3b är en upprepning av övningsuppgift 9.3a.
- s. 219: Notera att "Obs"-varningen för övningsuppgift 9.6c också gäller för övningsuppgift 9.6b: "Obs: Tänk på att en ökning i lägenhetsålder med ett år inte ökar lägenhetsålder i kvadrat med ett år."
- s. 234: Sista raden innan avsnitt 10.3.1 "I nästa avsnitt beskriver vi en *regulariseringsmetod*". Ska i stället vara "I nästa två avsnitt beskriver vi två *regulariseringsmetoder*".

- s. 239: Sista raden i

$D_1 = 1$ om $X_1 \geq 40$,	0 annars,
$D_2 = 1$ om $X_2 \geq 3\,000$,	0 annars,
$D_3 = 1$ om $X_1 \geq 100$,	0 annars,
$D_4 = 1$ om $X_2 \geq 3\,000$,	0 annars.

ska vara

$D_4 = 1$ om $X_2 \geq 8\,000$,	0 annars.
----------------------------------	-----------

- s. 261: I ekvation 11.6 ska det stå $E((Y_t - \mu_Y)(Y_{t-k} - \mu_Y))$ i stället för $E(Y_t - \mu_Y) E(Y_{t-k} - \mu_Y)$.
- s. 271: Variansskattningen i ekvation 11.30 är felaktig. Se bilaga A.5.2 för utförlig genomgång gällande detta.
- s. 274: Andra meningen efter ekvation 11.38. " $\hat{\delta}$ inte att följa en t -fördelning ..." ska vara " $\hat{\delta}/SE(\hat{\delta})$ inte att följa en t -fördelning ..."
- s. 279: Koden som används för att skapa månadsvariabeln i övningsuppgift 9.3c kräver paketet `lubridate`.
- s. 282: I fjärde stycket ska det stå "elever" i stället för "lever".
- s. 282: I sista stycket ska det stå "grupperna. Skälet" i stället för "grupperna.Skälet".
- s. 297: Raden innan ekvation 12.17. Inkludera "den förväntade" i meningen. Således "Något som är naturligt att vilja skatta är **den förväntade** skillnaden mellan behandling och kontroll i populationen, vilket vi skriver som

$$\tau^P = E(Y(1)) - E(Y(0)). \quad (1)$$

- s. 330: Rubriken till tabell 12.6 ska vara "Skattning av olika parametrar för olika bandbredder. Rad ett ger skattningar när den beroende variabeln är en binär variabel för vänstermajoritet, medan i rad två och tre är den beroende variabeln förändringar i skattesats. Den löpande variabeln är andel vänstermandat."
- s. 330: Tredje stycket tredje raden från slutet ska det stå "stickprovsstorleken ökar med bandbredden" i stället för "stickprovsstorleken minskar med bandbredden".

Referenser

- Deryugina, T., Heutel, G., Miller, N. H., Molitor, D., & Reif, J. (2019). The Mortality and Medical Costs of Air Pollution: Evidence from Changes in Wind Direction. *American Economic Review*, 109(12), 4178–4219.