Exam Formula Sheet

Epi 202: Probability

$$\begin{aligned}\operatorname{Var}\big(\tilde{a}\cdot\tilde{X}\big) &= \operatorname{Var}\left(\sum_{i=1}^{n} a_i X_i\right)\\ &= \tilde{a}^\top \operatorname{Var}\big(\tilde{X}\big)\tilde{a}\\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \operatorname{Cov}\big(X_i, X_j\big)\end{aligned}$$

$$\mathrm{E}[Y] = \mathrm{E}[\mathrm{E}[Y \mid X]]$$

$$\mathrm{E}[Y \mid Z] = \mathrm{E}[\mathrm{E}[Y \mid X, Z] \mid Z]$$

$$\operatorname{Var}(Y) = \mathrm{E}[\operatorname{Var}(Y \mid X)] + \operatorname{Var}(\mathrm{E}[Y \mid X])$$

$$\operatorname{Cov}(Y, Z) = \mathrm{E}[\operatorname{Cov}(Y, Z \mid X)] + \operatorname{Cov}(\mathrm{E}[Y \mid X], \mathrm{E}[Z \mid X])$$

Epi 203: Statistical inference

$$\mathcal{L}(\theta) \stackrel{\text{def}}{=} \mathrm{p}(\tilde{X} = \tilde{x} | \Theta = \theta)$$

$$\ell \stackrel{\text{def}}{=} \log\{\mathcal{L}(\tilde{x}|\theta)\}$$

$$\ell' \stackrel{\text{def}}{=} \frac{\partial}{\partial\theta}\ell(\tilde{x}|\theta)$$

$$\ell'' \stackrel{\text{def}}{=} \frac{\partial}{\partial\tilde{\theta}}\frac{\partial}{\partial\tilde{\theta}^\top}\ell(\tilde{x}|\tilde{\theta})$$

$$\ell''_{ij} = \frac{\partial}{\partial\theta_i}\frac{\partial}{\partial\theta_j}\ell(\tilde{X} = \tilde{x}|\tilde{\theta})$$

$$I \stackrel{\text{def}}{=} -\ell''(\tilde{x}|\tilde{\theta})$$

$$\mathcal{I} \stackrel{\text{def}}{=} \mathrm{E}[I(\tilde{x}|\theta)]$$

$$\hat{\theta}_{ML} \mathrel{\dot\sim} \mathrm{N}\Big(\theta, \big[\mathcal{I}(\tilde{\theta})\big]^{-1}\Big)$$

For one parameter θ_k:

$$\mathrm{CI}_{1-\alpha}(\theta_k) = \Big[\hat{\theta}_k \pm z_{1-\frac{\alpha}{2}}\,\widehat{\mathrm{SE}}\big(\hat{\theta}_k\big)\Big]$$

$$Z_k \stackrel{\text{def}}{=} \frac{\hat{\theta}_k - \theta_{k,0}}{\widehat{\mathrm{SE}}\big(\hat{\theta}_k\big)} \mathrel{\dot\sim} \mathrm{N}(0,1) \qquad \text{under } H_0 : \theta_k = \theta_{k,0} \qquad z_k = \text{observed } Z_k$$

$$\begin{aligned}p\text{-value} &= 2\Pr(|Z| \geq |z_k|), \quad Z \sim \mathrm{N}(0,1)\\ &= 2(1 - \Phi(|z_k|))\end{aligned}$$

Sta 108: Linear regression

$$t_k \stackrel{\text{def}}{=} \frac{\hat{\beta}_k - \beta_{k,0}}{\widehat{\text{SE}}\big(\hat{\beta}_k\big)} \mathrel{\dot\sim} t_{n-p} \qquad \text{under } H_0 : \beta_k = \beta_{k,0} \qquad t_k^{\text{obs}} = \text{observed } t_k$$

$$\text{CI}_{1-\alpha}(\beta_k) = \Big[\hat{\beta}_k \pm t_{n-p}\Big(1 - \frac{\alpha}{2}\Big)\,\widehat{\text{SE}}\big(\hat{\beta}_k\big)\Big]$$

Let $T_{n-p} \sim t_{n-p}$.

$$\begin{aligned} p\text{-value} &= 2\,\Pr\big(|T_{n-p}| \geq |t_k^{\text{obs}}|\big) \\ &= 2\Big(1 - F_{t_{n-p}}\big(|t_k^{\text{obs}}|\big)\Big) \end{aligned}$$

$$\text{CI}_{1-\alpha}(\mu(\tilde{x}^*)) = \Big[\hat{\mu}(\tilde{x}^*) \pm t_{n-p}\Big(1 - \frac{\alpha}{2}\Big)\,\widehat{\text{SE}}(\hat{\mu}(\tilde{x}^*))\Big]$$

Y^* denotes a new observation (not in the training data), with corresponding covariate pattern $\tilde{x}^*$. Let $\hat{Y}^* \stackrel{\text{def}}{=} \hat{\mu}(\tilde{x}^*)$.

$$\text{PI}_{1-\alpha}(Y^*|\tilde{x}^*) = \Big[\hat{Y}^* \pm t_{n-p}\Big(1 - \frac{\alpha}{2}\Big)\,\widehat{\text{SE}}\big(Y^* - \hat{Y}^*\big)\Big]$$

$$\widehat{\text{SE}}\big(Y^* - \hat{Y}^*\big) = \hat{\sigma}\sqrt{1 + (\tilde{x}^*)^\top (\mathbf{X}'\mathbf{X})^{-1}\tilde{x}^*}$$

$$\text{Var}\big(Y^* - \hat{Y}^*\big) = \sigma^2\big(1 + (\tilde{x}^*)^\top (\mathbf{X}'\mathbf{X})^{-1}\tilde{x}^*\big)$$

Let $\hat{\Sigma} \stackrel{\text{def}}{=} \widehat{\text{Var}}\big(\hat{\beta}\big) = \hat{\sigma}^2(\mathbf{X}'\mathbf{X})^{-1}$.

$$\widehat{\text{Var}}(\hat{\mu}(\tilde{x})) = \tilde{x}^\top \hat{\Sigma} \tilde{x}$$

Let $\Delta\mu(\tilde{x}, \tilde{x}^*) = \mu(\tilde{x}) - \mu(\tilde{x}^*)$, and let $\Delta\tilde{x} = \tilde{x} - \tilde{x}^*$; then:

$$\widehat{\text{Var}}\big(\widehat{\Delta\mu}(\tilde{x}, \tilde{x}^*)\big) = \Delta\tilde{x}^\top \hat{\Sigma} \Delta\tilde{x}$$

Epi 204: Generalized linear models

Generalized linear models have three components:

1. The **outcome distribution** family: $\text{p}(Y|\mu(\tilde{x}))$
2. The **link function**: $g(\mu(\tilde{x})) = \eta(\tilde{x})$
3. The **linear component**: $\eta(\tilde{x}) = \tilde{x} \cdot \beta$

$$\theta_\omega(\tilde{x}, \tilde{x}^*) = \exp\big\{(\Delta\tilde{x}) \cdot \tilde{\beta}\big\}$$

Estimates of odds ratios from 2x2 contingency tables

$$\hat{\theta} = \frac{ad}{bc}$$

Summary of logistic regression definitions and results

Odds and log-odds

Odds

$$\omega \stackrel{\text{def}}{=} \frac{\Pr(A)}{\Pr(\neg A)}$$

Conditional odds

$$\omega(A|B) \stackrel{\text{def}}{=} \frac{\Pr(A|B)}{\Pr(\neg A|B)}$$

Odds function

$$\text{odds}\{\pi\} \stackrel{\text{def}}{=} \frac{\pi}{1-\pi}$$

Probability to odds

$$\omega = \frac{\pi}{1-\pi}$$

Odds function equals odds

$$\omega = \text{odds}\{\pi\}$$

Simplified odds expressions

$$\text{odds}\{\pi\} = \frac{1}{\pi^{-1} - 1} = (\pi^{-1} - 1)^{-1}$$

Odds of a non-event

$$\omega(\neg A) = \frac{1 - \pi}{\pi} = \pi^{-1} - 1$$

Odds ratio

$$\theta(\omega_1, \omega_2) \stackrel{\text{def}}{=} \frac{\omega_1}{\omega_2}$$

OR as ratio of probability ratios

$$\begin{aligned}\theta(\omega_1, \omega_2) &= \frac{\omega_1}{\omega_2} \\ &= \frac{\left(\frac{\pi_1}{1-\pi_1}\right)}{\left(\frac{\pi_2}{1-\pi_2}\right)}\end{aligned}$$

Odds ratios are reversible

$$\theta_{\omega}(A|B) = \theta_{\omega}(B|A)$$

Conditional ORs are reversible

$$\theta_{\omega}(A|B, C) = \theta_{\omega}(B|A, C)$$

Inverse-odds and probability recovery

Odds to probability

$$\pi = \frac{\omega}{1 + \omega}$$

Inverse-odds function

$$\text{invodds}\{\omega\} \stackrel{\text{def}}{=} \frac{\omega}{1 + \omega}$$

Probability as inverse-odds

$$\pi = \text{invodds}\{\omega\}$$

Simplified inverse-odds

$$\text{invodds}\{\omega\} = \frac{1}{1 + \omega^{-1}} = (1 + \omega^{-1})^{-1}$$

One minus inverse-odds

$$1 - \pi = \frac{1}{1 + \omega}$$

Complement of inverse-odds

$$1 + \omega = \frac{1}{1 - \pi}$$

Log-odds (logit) and expit

Log-odds

$$\eta \stackrel{\text{def}}{=} \log\{\omega\}$$

Log-odds from probability

$$\eta = \log\Big\{\frac{\pi}{1 - \pi}\Big\}$$

Logit function

$$\text{logit}(\pi) \stackrel{\text{def}}{=} \log\{\text{odds}\{\pi\}\}$$

Logit expanded

$$\text{logit}(\pi) = \log\Big\{\frac{\pi}{1 - \pi}\Big\}$$

Log-odds equals logit

$$\eta = \text{logit}\{\pi\}$$

Odds from log-odds

$$\omega = \exp\{\eta\}$$

Probability from log-odds

$$\pi = \frac{\exp\{\eta\}}{1 + \exp\{\eta\}}$$

Expit / inverse-logit function

$$\text{expit}(\eta) \stackrel{\text{def}}{=} \text{invodds}\{\exp\{\eta\}\}$$

Expit expressions

$$\text{expit}(\eta) = \frac{\exp\{\eta\}}{1+\exp\{\eta\}} = (1+\exp\{-\eta\})^{-1}$$

Probability as expit

$$\pi = \text{expit}\{\eta\} \tag{1}$$

Logit and expit are inverses

$$\text{logit}\{\text{expit}\{\eta\}\} = \eta \qquad \text{expit}\{\text{logit}\{\pi\}\} = \pi$$

$$\overbrace{\Big[\pi \stackrel{\text{def}}{=} \Pr(Y=1|\tilde{X}=\tilde{x})\Big] \underbrace{\underset{\frac{\omega}{1+\omega}}{\overset{\frac{\pi}{1-\pi}}{\rightleftarrows}} \Big[\omega \stackrel{\text{def}}{=} \text{odds}(Y=1|\tilde{X}=\tilde{x})\Big] \underset{\exp\{\eta\}}{\overset{\log\{\omega\}}{\rightleftarrows}} \Big[\eta(\tilde{x}) \stackrel{\text{def}}{=} \text{log-odds}(Y=1|\tilde{X}=\tilde{x})\Big]}_{\text{expit}(\eta)}}^{\text{logit}(\pi)}$$

Figure 1: Diagram of logistic regression link and inverse link functions

Rare events

Odds minus probability

$$\omega - \pi = \frac{\pi^2}{1-\pi}, \quad \text{where } \omega = \frac{\pi}{1-\pi}$$

Derivatives

	π	ω	η	$\tilde{x}$	$\tilde{\beta}$
π	1	$(1+\omega)^2$	$\frac{(1+\omega)^2}{\omega}$	undef	undef
ω	$(1-\pi)^2$	1	$\frac{1}{\omega}$	undef	undef
η	$\pi(1-\pi)$	ω	1	undef	undef
$\tilde{x}$	$\tilde{\beta}\pi(1-\pi)$	$\tilde{\beta}\omega$	$\tilde{\beta}$	$\mathbb{I}$	$\mathbf{0}$
$\tilde{\beta}$	$\tilde{x}\pi(1-\pi)$	$\tilde{x}\omega$	$\tilde{x}$	$\mathbf{0}$	$\mathbb{I}$

Column labels indicate the numerators of the derivatives; row labels indicate the denominators.

Log-likelihood and score function

Log-likelihood component

$$\ell_i(\pi_i) = y_i\eta_i - \log\{1+\omega_i\}$$

Score function as sum

$$\tilde{\ell}'(\tilde{\beta}) = \sum_{i=1}^{n} \tilde{\ell}'_i(\tilde{\beta})$$

Score component

$$\ell'_i(\tilde{\beta}) = \tilde{x}_i e_i$$

Score function

$$\tilde{\ell}'(\tilde{\beta}) = \sum_{i=1}^{n} \tilde{x}_i e_i = \mathbf{X}^\top \tilde{e}$$

MLE

One-sample MLE for odds

$$\widehat{\omega} = \frac{x}{n-x}$$

Odds ratios in logistic regression

General OR formula

$$\theta_\omega(\tilde{x}, \tilde{x}^*) = \exp\{\eta(\tilde{x}) - \eta(\tilde{x}^*)\}$$

Difference in log-odds

$$\Delta\eta \stackrel{\text{def}}{=} \eta(\tilde{x}) - \eta(\tilde{x}^*)$$

OR in terms of $\Delta\eta$

$$\theta_\omega(\tilde{x}, \tilde{x}^*) = \exp\{\Delta\eta\}$$

$\Delta\eta$ from covariates

$$\Delta\eta = (\tilde{x} - \tilde{x}^*) \cdot \tilde{\beta}$$

Difference in covariate patterns

$$\Delta\tilde{x} \stackrel{\text{def}}{=} \tilde{x} - \tilde{x}^*$$

$\Delta\eta$ from $\Delta\tilde{x}$

$$\Delta\eta = \Delta\tilde{x} \cdot \tilde{\beta}$$

OR in terms of $\Delta\tilde{x}$

$$\theta_\omega(\tilde{x}, \tilde{x}^*) = \exp\Big\{(\Delta\tilde{x}) \cdot \tilde{\beta}\Big\}$$

Log OR equals $\Delta\eta$

$$\log\{\theta_\omega(\tilde{x}, \tilde{x}^*)\} = \Delta\eta$$

Inference for log-odds and odds ratios

Estimated SE of log-odds

$$\widehat{\mathrm{Var}}(\hat{\eta}(\tilde{x})) = \tilde{x}^\top \hat{\Sigma} \tilde{x}$$
$$\widehat{\mathrm{SE}}(\hat{\eta}(\tilde{x})) = \sqrt{\tilde{x}^\top \hat{\Sigma} \tilde{x}}$$

Estimated SE of $\Delta\hat{\eta}$

$$\widehat{\mathrm{Var}}(\Delta\hat{\eta}) = \Delta\tilde{x}^\top \hat{\Sigma} (\Delta\tilde{x})$$
$$\widehat{\mathrm{SE}}(\Delta\hat{\eta}) = \sqrt{\Delta\tilde{x}^\top \hat{\Sigma} (\Delta\tilde{x})}$$

Comparing probabilities

Risk difference

$$\delta(\pi_1, \pi_2) \stackrel{\text{def}}{=} \pi_1 - \pi_2$$

Risk ratio

$$\rho(\pi_1, \pi_2) \stackrel{\text{def}}{=} \frac{\pi_1}{\pi_2}$$

Relative risk difference

$$\xi(\pi_1, \pi_2) \stackrel{\text{def}}{=} \frac{\delta(\pi_1, \pi_2)}{\pi_2}$$

RRD equals RR minus 1

$$\xi(\pi_1, \pi_2) = \rho(\pi_1, \pi_2) - 1$$

Logistic regression model

Logistic regression

$$Y_i | \tilde{X}_i \ \sim_{\perp\!\!\!\perp} \ \mathrm{Ber}(\pi(\tilde{X}_i)), \qquad \mathrm{logit}\{\pi(\tilde{x})\} = \tilde{x}^\top \tilde{\beta}$$

Survival analysis

Probability distribution functions

Table 1: Probability distribution functions

Name	Symbols	Definition
Probability density function (PDF)	$\mathrm{f}(t), \mathrm{p}(t)$	$\mathrm{p}(T = t)$
Cumulative distribution function (CDF)	$\mathrm{F}(t), \mathrm{P}(t)$	$\mathrm{P}(T \leq t)$
Survival function	$\mathrm{S}(t), \bar{\mathrm{F}}(t)$	$\mathrm{P}(T > t)$
Hazard function	$\lambda(t), \mathrm{h}(t)$	$\mathrm{p}(T = t \vert T \geq t)$
Cumulative hazard function	$\Lambda(t), \mathrm{H}(t)$	$\int_{u=-\infty}^{t} \lambda(u)du$
Log-hazard function	$\eta(t)$	$\log\{\lambda(t)\}$

Diagram of survival distribution function relationships

$$\mathrm{f}(t) \xleftarrow[\mathrm{S}(t)\lambda(t)]{-S'(t)} \mathrm{S}(t) \xleftarrow{\exp\{-\Lambda(t)\}} \Lambda(t) \xleftarrow{\int_{u=0}^{t}\lambda(u)du} \lambda(t) \xleftarrow{\exp\{\eta(t)\}} \eta(t)$$

$$\mathrm{f}(t) \xrightarrow[\int_{u=t}^{\infty}\mathrm{f}(u)du]{\mathrm{f}(t)/\lambda(t)} \mathrm{S}(t) \xrightarrow[-\log \mathrm{S}(t)]{} \Lambda(t) \xrightarrow[\Lambda'(t)]{} \lambda(t) \xrightarrow[\log\{\lambda(t)\}]{} \eta(t)$$

Survival likelihood contributions, assuming non-informative censoring

$$\begin{aligned}\mathrm{p}(Y = y, D = d) &= [\mathrm{f}_T(y)]^d[\mathrm{S}_T(y)]^{1-d} \\ &= [\lambda_T(y)]^d[\mathrm{S}_T(y)]\end{aligned}$$

Nonparametric time-to-event distribution estimators

$$\hat{\lambda}_i = \frac{d_i}{r_i}$$

$$\hat{\kappa}_i = 1 - \hat{\lambda}_i = \frac{r_i - d_i}{r_i}$$

$$\hat{\mathrm{S}}_{KM}(t) \stackrel{\text{def}}{=} \prod_{\{i:\ t_i \leq t\}} \hat{\kappa}_i$$

$$\hat{\Lambda}_{NA}(t) \stackrel{\text{def}}{=} \sum_{\{i:\ t_i \leq t\}} \hat{\lambda}_i$$

Proportional hazards model structure

Formula	Description
$\lambda(t \mid \tilde{x}) = \lambda_0(t) \cdot \theta_\lambda(\tilde{x})$	Proportional hazards assumption
$\Lambda(t \mid \tilde{x}) = \Lambda_0(t) \cdot \theta_\lambda(\tilde{x})$	Cumulative hazard factorization
$\eta(t \mid \tilde{x}) = \eta_0(t) + \Delta\eta(\tilde{x})$	Log-hazard decomposition
$\Delta\eta(\tilde{x}) = \tilde{x} \cdot \tilde{\beta} = \beta_1 x_1 + \cdots + \beta_p x_p$	Linear predictor
$\theta_\lambda(\tilde{x}) = \exp\{\Delta\eta(\tilde{x})\}$	Hazard multiplier
$\theta_\lambda(t \vert \tilde{x} : \tilde{x}^*) \stackrel{\text{def}}{=} \frac{\lambda(t\vert\tilde{x})}{\lambda(t\vert\tilde{x}^*)}$	Hazard ratio definition
$\theta_\lambda(t \mid \tilde{x} : \tilde{x}^*) = \exp\{\Delta\eta(\tilde{x}) - \Delta\eta(\tilde{x}^*)\} = \exp\Big\{(\tilde{x} - \tilde{x}^*) \cdot \tilde{\beta}\Big\}$	Hazard ratio formula

Proportional hazards model partial likelihood formula:

$$\mathcal{L}_i^* = \frac{\theta_\lambda(\tilde{x}_i)}{\sum_{k \in R(t_i)} \theta_\lambda(\tilde{x}_k)}$$

$$\mathcal{L}^* = \prod_{\{i:\ d_i = 1\}} \mathcal{L}_i^*$$

Proportional hazards model baseline cumulative hazard estimator:

$$\hat{\Lambda}_0(t) = \sum_{t_i < t} \frac{d_i}{\sum_{k \in R(t_i)} \theta_\lambda(x_k)}$$