1. Combinatorics

repetition and ordered? $\to n^r$
no repetition and ordered? $\to P(n,r) = \frac{n!}{(n-r)!}$
no repetition and not ordered? $\to C(n,r) = \frac{n!}{r!(n-r)!}$
repetition and not ordered? $\to C(n+r-1,r) = \frac{(n+r-1)!}{r!(n-1)!}$

2. Probability

Prob. space: Ω: realisations, $P : \mathcal{P}(\Omega) \to [0,1]$:
$\bullet P(A\cup B) = P(A) + P(B) - P(A\cap B)$
$A_i \cap A_j = \varnothing \Rightarrow P(\bigcup A_i) = \sum P(A_i)$ (finite/countable)
$P(\bigcup A_i) \leq \sum P(A_i)$ (σ-sub-additivity);$A \subset B \Rightarrow P(A) \leq P(B)$
$A_i \subset A_{i+1} \Rightarrow \lim\limits_{n\to\infty} P(A_n) = P(\bigcup A_i)$; $A_i \supset A_{i+1} \Rightarrow \lim\limits_{n\to\infty} P(A_n) = P(\bigcap A_n)$
pmf: Ω finite/countable, $p : \Omega \to [0,1]$, $\sum_\omega p(\omega) = 1$
$P_p(A) = \sum_{\omega\in A} p(\omega)$ (probab. measure)
pdf: $f : \mathbb{R}^d \to [0,+\infty)$, $\int\cdots\int f(x_1,\ldots,x_d)dx = 1$
$P_f(A) = \int\cdots\int f(x)\mathbf{1}_A(x)dx$ (probab. measure)
$f_X(x) = \int_{\mathbb{R}} f_{X,Y}(x,y)dy$
$f_Y(y) = \left|\frac{dg^{-1}(y)}{dy}\right| f_X(g^{-1}(y))$ with $Y = g(X)$, g monotone
Def 2.2.3 (Continuous r.v.): $X : \Omega \to \mathbb{R}$ continuous if $\exists f_X : \mathbb{R} \to [0,+\infty)$: $P(X\in A) = \int_{\mathbb{R}} \mathbf{1}_A(x) f_X(x)dx$
Def 2.2.5 (CDF): $F_X(t) = P(X\leq t)$,X contin. $\Rightarrow F'_X(t) = f_X(t)$
CDF: right-cont., non-decr., $\lim_{t\to-\infty} F = 0$, $\lim_{t\to+\infty} F = 1$
Def 2.2.8 (Expectation):
Discrete: $\sum\limits_{w\in\Omega} |X(\omega)|p(\omega) < \infty \Rightarrow E_p(X) = \sum\limits_{w\in\Omega} X(\omega)p(\omega)$
Continuous: $\int_{\mathbb{R}^d} |X(x)|f(x)dx < \infty \Rightarrow E_f(X) = \int_{\mathbb{R}^d} X(x)f(x)dx$
Trnsfr: $E(g(X)) = \sum\limits_{x\in Image(X)} g(x)P(X=x)$ or $\int_{-\infty}^{+\infty} f_X(x)g(x)dx$
(Random vector): $X : \Omega \to \mathbb{R}^d$, $X(\omega) = (X_1(\omega),\ldots,X_d(\omega))$
$F_X(t_1,\ldots,t_d) = P(X_1\leq t_1,\ldots,X_d\leq t_d)$
Def 2.2.11: X discrete: $\exists\mathcal{D}_X$ countable, $P(X\in\mathcal{D}_X) = 1$
X continuous: $\exists f_X$: $\mathbb{R}^d \to [0,+\infty)$,P($X\in A$) $= \int_{\mathbb{R}^d} \mathbf{1}_A(x)f_X(x)dx$
Def 2.2.12 (Moments): $E(|X|^p) < \infty \Rightarrow p$-th moment: $E(X^p)$
Def 2.2.13 (MGF): $E(e^{\delta|X|}) < \infty \Rightarrow M_X(t) = E(e^{tX})$, $t\in(-\delta,\delta)$
Thm 2.2.5: $E(e^{\delta|X|}), E(e^{\delta|Y|}) < \infty \Rightarrow$
$\bullet$ X has moments of all orders; $M_X^{(n)}(0) = E(X^n)$
$\bullet$ $M_X(t) = M_Y(t)\ \forall t\in(-\delta,\delta) \Rightarrow X \overset{Law}{=} Y$ $P(A) > 0$: $P(B|A) = \frac{P(A\cap B)}{P(A)}$; $E_P(X|A) = E_{P_A}(X) = \frac{E_P(X\mathbf{1}_A)}{P(A)}$
Def 2.4.2: A, B indep. $\Leftrightarrow P(A\cap B) = P(A)P(B)$
$(A_i)_{i\in I}$ indep. family: $\forall J\subset I$ finite, $P(\bigcap_{i\in J} A_i) = \prod_{i\in J} P(A_i)$
Def 2.4.3 (Indep. r.v.): X, Y indep. $\Leftrightarrow \forall A, B\subset\mathbb{R}$: $P(X\in A, Y\in B) = P(X\in A)P(Y\in B)$
$\Leftrightarrow \forall f, g : \mathbb{R}\to\mathbb{R}$: $E(f(X)g(Y)) = E(f(X))E(g(Y))$
$(X_i)_{i\in I}$ indep. family: $\forall J\subset I$ finite, $\forall A_i\subset\mathbb{R}$, $i\in J$:
$P(\cap_{i\in J}\{X_i\in A_i\}) = \prod_{i\in J} P(X_i\in A_i)$
$\Leftrightarrow \forall J\subset I$ finite, $\forall f_i : \mathbb{R}\to\mathbb{R}$, $i\in J$:
$E(\prod_{i\in J} f_i(X_i)) = \prod_{i\in J} E(f_i(X_i))$
Def 2.4.4 : i.i.d. $\Leftrightarrow (X_i)_{i\in I}$ indep. family & $\forall i,j\in I$: $X_i \overset{Law}{=} X_j$
Thm 2.4.1: $d, d'\geq 1$. $X : \Omega\to\mathbb{R}^d$, $Y : \Omega\to\mathbb{R}^{d'}$ r.v.
$\bullet$ X, Y cont.: indep. $\Leftrightarrow f_{(X,Y)}(x,y) = f_X(x)f_Y(y)$
$\bullet$ X, Y discrete: indep. $\Leftrightarrow \forall x\in\mathbb{R}^d, y\in\mathbb{R}^{d'}$: $P(X=x, Y=y) = P(X=x)P(Y=y)$
$\bullet$ X discrete, Y cont.: indep. $\Leftrightarrow \forall x\in\mathbb{R}^d, A\subset\mathbb{R}^{d'}$:
$P(X=x, Y\in A) = P(X=x)\int_A f_Y(y)dy$
Lem 2.4.2 (Bayes): $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$
Total prob.: (A_i) partition, $P(A_i) > 0$:
$P(B) = \sum\limits_i P(B|A_i)P(A_i)$; $E(X) = \sum\limits_i E(X|A_i)P(A_i)$
Var$_P(X) = E_P((X - E_P(X))^2) = E_P(X^2) - E_P(X)^2$
Cov$_P(X,Y) = E_P(XY) - E_P(X)E_P(Y)$; $= 0 \Rightarrow$ uncorrelated
Lem 2.5.2: Cov(X,Y) = Cov(Y,X); Cov$(aX, bY) = ab$Cov(X,Y)
Cov$(X, Y_1 + Y_2)$ = Cov(X, Y_1) + Cov(X, Y_2)
Def 2.5.4 (Pearson): $\rho_{XY} = \frac{\text{Cov}(X,Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} = \frac{\text{Cov}(X,Y)}{\sigma_X\sigma_Y} \in [-1,1]$
$|\rho_{XY}| = 1 \Leftrightarrow \exists a, b\in\mathbb{R}$: $Y = aX + b$
Lem 2.5.3: $\min_{a,b\in\mathbb{R}} E((Y - aX - b)^2) = (1-\rho_{XY}^2)Var(Y)$
attained at $a = \frac{\text{Cov}(X,Y)}{\text{Var}(X)}$, $b = E(Y - aX)$

2.6 Classical Distributions TODO Add proba of laws

$X_1,\ldots,X_n \sim$ Bern(p) indep., $Y = \sum X_k \Rightarrow Y \sim$ Bin(n,p)
$X_1,\ldots \sim$ Bern(p) i.i.d., $Y = 1 + \sum\limits_{n\geq1} \prod_{i=1}^n (1 - X_i) \Rightarrow Y \sim$ Geo(p)
(Memory loss):$X \sim$ Geo(p): $P(X=n|X>k) = P(X = n-k)$
Under $P(\cdot|X>k)$, $X - k \sim$ Geo(p)
$X \sim$ Poi$(\lambda) \Leftrightarrow$ [P(X=0) = e$^{-\lambda}$ and $\forall k\in\mathbb{N}$: $\frac{P(X=k+1)}{P(X=k)} = \frac{\lambda}{k+1}$]
Lem 2.6.7: X, Y indep. Gaussian, $X\sim\mathcal{N}(\mu_1,\sigma_1^2)$, $Y\sim\mathcal{N}(\mu_2,\sigma_2^2)$:
1. $\tilde{X} = (X-\mu_1)/\sigma_1 \sim \mathcal{N}(0,1)$
2. $Z = X + Y \sim \mathcal{N}(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$
$\lambda > 0$, $X\sim$ Exp(λ). $\forall 0 < a < b$: $P(X\geq b|X\geq a) = P(X\geq b-a)$
Under $P(\cdot|X>a)$, $X - a \sim$ Exp(λ)

2.6.3 MGF

Bern(p): $1 + p(e^t - 1)$
Bin(n,p): $(1 + p(e^t-1))^n$ — Geo(p): $\frac{pe^t}{1-(1-p)e^t}$
Pois(λ): $\exp(\lambda(e^t - 1))$ — Uni$(\{0,\ldots,n\})$: $\frac{e^{(n+1)t}-1}{(n+1)(e^t-1)}$
Uni$(\{1,\ldots,n\})$: $\frac{e^{nt}-1}{n(1-e^{-t})}$ — Uni([a,b]):$\frac{e^{bt}-e^{at}}{t(b-a)}$
$\mathcal{N}(\mu,\sigma^2)$: $\exp(\mu t + \frac{\sigma^2t^2}{2})$ — Exp(λ): $\frac{\lambda}{\lambda-t}$

2.7 Inequalities and wLLN

Markov's inequality: $X\geq 0 \Rightarrow P(X\geq a) \leq \frac{E(X)}{a}$, $\forall a > 0$
Chebychev: g increasing, $g(X)$ r.v.: $P(X\geq a) \leq \frac{E(g(X))}{g(a)}$
p-th moment: $P(X\geq a) \leq \frac{E(|X|^p)}{a^p}$, $\forall a > 0, p\in(0,+\infty)$
Exponential: $P(X\geq a) \leq e^{-\delta a}E(e^{\delta X})$, $\forall a > 0$, $\delta\in\mathbb{R}_+$
wLLN, 2nd moment vers.: $X_1, X_2,\ldots$ identically distributed, $E(X_1^2) < \infty$, Cov$(X_i,X_j) = 0$ if $i\neq j$, $\bar{S}_n = \frac{1}{n}\sum X_i$:
$P(|\bar{S}_n - \mu|\geq\varepsilon) \leq \frac{\sigma^2}{\varepsilon^2 n} \Rightarrow \bar{S}_n \xrightarrow{proba} \mu$, ($\mu = E(X_1)$, $Var(X_1) = \sigma^2$)
$\bullet$E(XY)$^2 \leq E(X^2)E(Y^2)$ $\bullet$|Cov(X,Y)|$\leq \sqrt{\text{Var}(X)}\sqrt{\text{Var}(Y)}$
Jensen: $I\subset\mathbb{R}$, $g : I\to\mathbb{R}$ convex, $X : \Omega\to I$ r.v. ($P(X\in I) = 1$): $E(g(X)) \geq g(E(X))$. For $h : I\to\mathbb{R}$ concave: $E(h(X)) \leq h(E(X))$

3. Limit Theorems

Law: $X, X_1, X_2,\ldots : \Omega\to\mathbb{R}$ r.v. $X_n \xrightarrow{\text{Law}} X \Leftrightarrow$:
1.F$_{X_n}(t) \to F_X(t)$ $\forall t\in\mathbb{R}$: F_X cont. at t
2. $E_P(\varphi(X_n)) \xrightarrow{n\to\infty} E_P(\varphi(X))$ $\forall\varphi : \mathbb{R}\to\mathbb{R}$ cont. & bounded
Law: $X, X_1,\ldots : \Omega\to\mathbb{R}$ r.v. If $\exists\delta > 0$: $\sup_{|t|<\delta} E(e^{tX}) < +\infty$, $\sup_{|t|<\delta} E(e^{tX_i}) < +\infty$ $\forall i\geq 1$, $E(e^{tX_n}) \xrightarrow{n\to\infty} E(e^{tX})$ $\forall |t| < \delta$
$\Rightarrow$ X$_n \xrightarrow{\text{Law}} X$ as $n\to\infty$ and $\forall p\geq 1$: $E(X_n^p) \xrightarrow{n\to\infty} E(X^p)$
Law: $X : \Omega\to\mathbb{R}$ r.v. with Image$(X)\subset\mathbb{Z}$ **!**, $X_1, X_2,\ldots : \Omega\to\mathbb{R}$ r.v. $X_n \xrightarrow{\text{Law}} X$ iff $\forall k\in\mathbb{Z}$: $\lim\limits_{\epsilon\downarrow0}\lim\limits_{n\to\infty} P(X_n\in(k-\epsilon, k+\epsilon)) = P(X = k)$
Proba: $X, X_1,\ldots : \Omega\to\mathbb{R}$ r.v. $X_n \xrightarrow{\text{Proba}} X \Leftrightarrow \forall\epsilon > 0$: $\lim_{n\to\infty} P(|X - X_n|\geq\epsilon) = 0$
a.s: $X, X_1,\ldots : \Omega\to\mathbb{R}$ r.v. $X_n \xrightarrow{a.s.} X \Leftrightarrow$ P$(\lim_{n\to\infty} X_n = X) = 1$
Lem 3.1.3: $P(\lim_{n\to\infty} X_n = X) = 1 \Leftrightarrow$
$\forall\epsilon > 0$, $\lim_{n\to\infty} P(\sup_{m\geq n}|X_m - X|\geq\epsilon) = 0$
L^p $: X_n \xrightarrow{L^p} X$: $\lim_{n\to\infty} E[|X_n - X|^p] = 0$
$\bullet$ L$^\infty \Rightarrow$ a.s. $\Rightarrow$ Proba $\Rightarrow$ Law
$\bullet$ L$^\infty \Rightarrow L^p \Rightarrow$ Proba $\Rightarrow$ Law
wLLN: $X, X_1,\ldots$ indep. fam. of id. distr. r.v.
If $E(|X|) < \infty$, $\forall\epsilon > 0$: $\lim_{n\to\infty} P(|\frac{1}{n}\sum_{i=1}^n X_i - E(X)|\geq\epsilon) = 0$
I.e. $\frac{1}{n}\sum_{i=1}^n X_i \xrightarrow{\text{Proba}} E(X)$ as $n\to\infty$
Quant. wLLN: $X, X_1,\ldots$ indep. fam. of id. distr. r.v.
If $E(X^2) < \infty$, $\forall\epsilon > 0$, $\forall n\geq 1$: $P(|\frac{1}{n}\sum_{i=1}^n X_i - E(X)|\geq\epsilon) \leq \frac{\text{Var}(X)}{\epsilon^2 n}$
sLLN: $X, X_1,\ldots$ indep. fam. of id. distr. r.v.
If $E(|X|) < \infty$: $\frac{1}{n}\sum_{i=1}^n X_i \xrightarrow{a.s.} E(X)$ as $n\to\infty$
CLT: $X, X_1,\ldots$ indep. fam. of id. distr. r.v.
If $E(X^2) < \infty$: $\frac{1}{\sqrt{\sigma_X^2 n}}\sum_{i=1}^n (X_i - \mu) \xrightarrow{\text{Law}} \mathcal{N}(0,1)$ as $n\to\infty$

(Berry-Esseen CLT): $X, X_1,\ldots$ indep. fam. of id. distr. r.v.
If $E(|X|^3) < \infty$, $\forall n\geq 1$,(Z $\sim\mathcal{N}(0,1)$, $\mu = E(X)$, σ_X^2 = Var(X)) :
$\sup_{t\in\mathbb{R}} |P(\frac{1}{\sqrt{\sigma_X^2 n}}\sum_{i=1}^n (X_i - \mu)\leq t) - P(Z\leq t)| \leq \frac{0.5E(|X|^3)}{\sqrt{n}}$
$\bullet$ X,X$_1,\ldots$ indep. fam. of id. distr. r.v. If $\exists\delta > 0$: $E(e^{\delta|X|}) < \infty$:
$\frac{1}{\sqrt{\sigma_X^2 n}}\sum_{i=1}^n (X_i - \mu) \xrightarrow{\text{Law}} \mathcal{N}(0,1)$ as $n\to\infty$ ($\mu = E(X)$, ...)

4. Statistical Inference

• For $n \geq 1$, $X_1, \dots, X_n$ n-sample of law $\mathbb{P}_\theta$, $\hat{f}_n$ estimator of (θ).
$(\hat{f}_n)_{n\geq 1}$ is **convergent** (or **consistent**) if $\hat{f}_n \xrightarrow{\text{Proba}}$ f(θ):
$\forall \epsilon > 0$, $\forall \theta \in \Theta$: $\lim_{n\to\infty} P(|\hat{f}_n(X_1, \dots, X_n) - f(\theta)| \geq \epsilon) = 0$

• For $n \geq 1$, $X_1, \dots, X_n$ n-sample of law $\mathbb{P}_\theta$, $\hat{f}$ estimator of $f(\theta)$.
Bias$_\theta(\hat{f}) = E(\hat{f}(X_1, \dots, X_n)) - f(\theta)$
$\hat{f}$ is *unbiased* if $\forall \theta \in \Theta$: Bias$_\theta(\hat{f}) = 0$. Otherwise *biased*

• For $n \geq 1$, $X_1, \dots, X_n$ n-sample of law $\mathbb{P}_\theta$, $\hat{f}_n$ estimator of $f(\theta)$.
$(\hat{f}_n)_{n\geq 1}$ is *asymptotically unbiased* if $\forall \theta \in \Theta$: $\lim_{n\to\infty}$ Bias$_\theta(\hat{f}_n) = 0$

• For $n \geq 1$, $X_1, \dots$ n-sample of law $\mathbb{P}_\theta$, $\hat{f}_n$ estimator of $f(\theta)$.
$\hat{f}_n$ is asymptotically unbiased **&** Var$_\theta(\hat{f}_n) \xrightarrow{n\to\infty} 0$
$\Rightarrow \hat{f}_n$ is a **convergent sequence of estimators**

4.1.3 Example of estimators

Emp. mean: $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ (!bias.); converg. if $E(|X_1|) < \infty$

Emp. median: $\tilde{X}_n = \begin{cases} \tilde{X}_{\frac{n+1}{2}} & \text{if } n \text{ odd} \\ \frac{1}{2}(\tilde{X}_{\frac{n}{2}} + \tilde{X}_{1+\frac{n}{2}}) & \text{if } n \text{ even} \end{cases}$ $(\tilde{X}_{i+1} \geq \tilde{X}_i)$

$P(X \leq M) = P(X \geq M) = \frac{1}{2}$

Emp. variance: $\hat{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X}_n)^2 = \overline{X^2}_n - \bar{X}_n^2$ (bias.)
$S_n^2 = \frac{n}{n-1}\hat{\sigma}_n^2$ (!bias.); converg. if $E(X_1^2) < \infty$

Emp. covariance: $\hat{\tau}_n = \frac{1}{n}\sum_{i=1}^n X_iY_i - \bar{X}_n\bar{Y}_n$ (bias.); $\frac{n}{n-1}\hat{\tau}_n$ (!bias.)

4.2 Method of Moments

Moment method: Est. θ from n-sample $X_1, \dots, X_n$ of law $\mathbb{P}_\theta$:
1. Find h, g: $\theta = h(E(g(X)))$ with $X \sim \mathbb{P}_\theta$
2. Use emp. mean $\frac{1}{n}\sum_{i=1}^n g(X_i)$ to est. $E(g(X))$
3. Estimator: $\hat{\theta}_n = h\left(\frac{1}{n}\sum_{i=1}^n g(X_i)\right)$

4.3 MLE

Def 4.2.1 (Likelihood): $n \geq 1$, sample $X_1, \dots, X_n$ of law $\mathbb{P}_\theta$. For realisation $x_1, \dots, x_n$, likelihood $\mathcal{L}(x_1, \dots, x_n|\theta)$:
• $\mathbb{P}_\theta$ discrete: $\mathcal{L}(x_1, \dots, x_n|\theta) = P(X_1 = x_1, \dots, X_n = x_n) = \prod_{i=1}^n P(X_i = x_i) = \prod_{i=1}^n \mathbb{P}_\theta(\{x_i\})$
• $\mathbb{P}_\theta$ continuous: $\mathcal{L}(x_1, \dots, x_n|\theta) = \prod_{i=1}^n f_{X_i}(x_i) = \prod_{i=1}^n f_\theta(x_i)$

Def 4.2.2 (MLE): Maximum likelihood estimator:
MLE$(X_1, \dots, X_n) = \arg\max_\theta \mathcal{L}(X_1, \dots, X_n|\theta)$

min/max : $f'(\theta^*) = 0$ (point critique) $f''(\theta^*) > 0 \Rightarrow$ min. loc.; $f''(\theta^*) < 0 \Rightarrow$ max. loc.; $f''(\theta^*) = 0 \Rightarrow$ non concluant

4.4 Confidence Intervals

(q-quantile): X r.v., $q > 0$ integer. $t \in \mathbb{R}$ is kth q-quantile of X $(k \geq 1)$ if: $P(X < t) \leq \frac{k}{q}$ AND $P(X \leq t) \geq \frac{k}{q}$
X cont. r.v. with strictly pos. density: only one kth q-quantile.
ν-quantiles with $\nu \in (0, 1)$: kth ν-quantile is $t \in \mathbb{R}$: $P(X < t) \leq k\nu$ AND $P(X \leq t) \geq k\nu$

Def: $X_1, \dots, X_n$ sample of law $\mathbb{P}_\theta$. $\alpha \in (0, 1)$. Random interval $I = I(X_1, \dots, X_n)$ not depending on θ is level $1 - \alpha$ **confidence interval** for $f(\theta)$ if $\forall \theta \in \Theta$: $P(f(\theta) \in I(X_1, \dots, X_n)) = 1 - \alpha$

Def: $I = I(X_1, \dots, X_n)$ is excess **confidence interval** for $f(\theta)$ at level $1 - \alpha$ if $P(f(\theta) \in I(X_1, \dots, X_n)) \geq 1 - \alpha$

4.5 Hypothesis Testing

4.5.1 General principle

Def 4.5.1: "$\theta \in \Theta_0$" (usually H_0): null hypothesis
"$\theta \in \Theta \setminus \Theta_0$" (usually H_1): alternative hypothesis

Rejection region: D event for r.v. $X_1, \dots, X_n$. I.e.: if X_i take values in $\mathbb{R}^d$, $D \subset (\mathbb{R}^d)^n$. In practice:
$D = \{(x_1, \dots, x_n) : T(x_1, \dots, x_n) \in [a, b]\}$ for a statistic T, $a < b$

Test procedure: Given D rejection region, H_0, H_1 hypotheses:
1. reject H_0 if $(X_1, \dots, X_n) \in D$
2. do not reject H_0 if $(X_1, \dots, X_n) \notin D$

Failure types:
• Type-I error: reject H_0 whereas correct
• Type-II error: do not reject H_0 whereas false
• $\alpha \in [0, 1]$. Test has **risk level** α or **confidence level** $1 - \alpha$ if $\sup_{\theta\in\Theta_0} P((X_1, \dots, X_n) \in D) = \alpha$

Power of test: $\inf_{\theta\in\Theta_1} P((X_1, \dots, X_n) \in D) = 1 - \beta$
In particular: $\beta = \sup_{\theta\in\Theta_1} P((X_1, \dots, X_n) \notin D)$

4.5.2 χ^2 test (adequacy)

H_0: sample from law $\mathbb{Q}$. Partition $\Omega = \bigcup_{i=1}^k \Omega_i$, $q_i = \mathbb{Q}(\Omega_i)$
$N_{n,i} = \sum_{j=1}^n \mathbf{1}_{\Omega_i}(X_j)$ (observed freq.), $Z_n = \sum_{i=1}^k \frac{(N_{n,i} - nq_i)^2}{nq_i}$
Under H_0: $Z_n \xrightarrow{\text{Law}} \chi^2_{k-1}$; Under H_1: $Z_n \xrightarrow{a.s.} +\infty$
Rejection: $D = \{Z_n > C\}$, C: $P(\chi^2_{k-1} > C) = \alpha$

4.5.3 t-test

$Z \sim \mathcal{N}(0, 1)$, $V \sim \chi^2_k$ indep. $\Rightarrow Z\sqrt{\frac{k}{V}} \sim$ Student$_t(k)$

One-sample: $X_i \sim \mathcal{N}(\mu, \sigma^2)$, H_0: "$\mu = E(X_1) = \mu_0$", $\mu_0 \in \mathbb{R}$
$T = \frac{\sqrt{n}(\bar{X}_n - \mu_0)}{\sqrt{\frac{1}{n-1}\sum(X_i - \bar{X}_n)^2}} \sim$ Student$_t(n-1)$ under H_0

Two-sample: $X_i \sim \mathcal{N}(\mu_X, \sigma^2)$, $Y_i \sim \mathcal{N}(\mu_Y, \sigma^2)$ indep.
$T_n = \frac{\sqrt{n}(\bar{X}_n - \bar{Y}_n)}{\sqrt{s_X^2 + s_Y^2}} \sim$ Student$_t(2n-2)$ under $H_0 =$ "$\mu_X = \mu_Y$"
$\rightarrow$Rejection: $\{|T| \geq C\}$ or $\{|T_n| \geq C\}$, $C > 0$ from confidence level

4.5.4 F-test (variance comparison)

$X \sim \chi^2_{k_1}$, $Y \sim \chi^2_{k_2}$ indep. $\Rightarrow \frac{X/k_1}{Y/k_2} \sim$ F(k_1, k_2)

$X \sim$ F(n, m): $E(X) = \frac{m}{m-2}$ $(m > 2)$, Var$(X) = \frac{2m^2(m+n-2)}{n(m-2)^2(m-4)}$ $(m > 4)$

Equality of var: $X_i \sim \mathcal{N}(\mu_X, \sigma_X^2)$, $Y_j \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$ indep.
$S_X^2 = \frac{1}{n-1}\sum(X_i - \bar{X})^2$, $S_Y^2 = \frac{1}{m-1}\sum(Y_j - \bar{Y})^2$
$\frac{S_X^2}{S_Y^2} \overset{\text{Law}}{=} \frac{\sigma_X^2}{\sigma_Y^2} \cdot Z$, $Z \sim$ F$(n-1, m-1)$.
$P(Z \leq 1 + c_\alpha^+) = 1 - \frac{\alpha}{2}$, $\quad P(Z \leq 1 - c_\alpha^-) = \frac{\alpha}{2}$
Under H_0 $(\sigma_X^2 = \sigma_Y^2)$: $\frac{S_X^2}{S_Y^2} \sim$ F$(n-1, m-1)$
Rejection: $D = \left\{\frac{S_X^2}{S_Y^2} \notin [1 - c_\alpha^-, 1 + c_\alpha^+]\right\}$

ANOVA (one-way)

k groups, $X_{l,i} \sim \mathcal{N}(\mu_l, \sigma^2)$ indep., n_l samples/group, $N = \sum n_l$
H_0: $\bar{\mu}_l = \bar{\mu}\ \forall l$ (group means = global mean)
$\overline{X}_l = \frac{1}{n_l}\sum_{i=1}^{n_l} X_{l,i}$, $\overline{X} = \frac{1}{N}\sum_{l,i} X_{l,i}$
$Y_{\text{bet}} = \frac{1}{k-1}\sum_{l=1}^k n_l(\overline{X}_l - \overline{X})^2$, $Y_{\text{in}} = \frac{1}{N-k}\sum_{l=1}^k\sum_{i=1}^{n_l}(X_{l,i} - \overline{X}_l)^2$
$Z = \frac{Y_{\text{bet}}}{Y_{\text{in}}} \sim$ F$(k-1, N-k)$ under H_0; $Z \to \infty$ under H_1
Rejection: $D = \{Z > c\}$, $P($F$(k-1, N-k) > c) = \alpha$

4.6 Comparing estimators

4.6.1 Mean square error

• n ≥ 1, $X_1, \dots, X_n$ n-sample of law $\mathbb{P}_\theta$, $\hat{f}$ estimator of $f(\theta)$.
Mean square error: MSE$_\theta(\hat{f}) = E((\hat{f}(X_1, \dots, X_n) - f(\theta))^2)$
• n ≥ 1, $X_1, \dots, X_n$ n-sample of law $\mathbb{P}_\theta$, $f(\theta)$ quantity to estimate.
Minimal mean square error estimator of $f(\theta)$:
MMSE$_{f(\theta)}(X_1, \dots, X_n) = \arg\min_{\hat{f}} E((\hat{f}(X_1, \dots, X_n) - f(\theta))^2)$
, min. over all estimators $\hat{f}(X_1, \dots, X_n)$ of $f(\theta)$. Existence non-trivial, generally not easy to compute

4.6.2 Asymptotic normality

• n ≥ 1, $X_1, \dots, X_n$ n-sample of law $\mathbb{P}_\theta$, $\hat{f}_n$ estimator of $f(\theta)$. $\hat{f}_n$ is **asymptotically normal sequence of estimators** if:
1. $\hat{f}_n$ is convergent
2. $\exists C \geq 0$: $\sqrt{n}(\hat{f}_n - f(\theta)) \xrightarrow{\text{Law}} \mathcal{N}(0, C)$ as $n \to \infty$

Other

e$^x = \sum_{k=0}^n \frac{x^k}{k!} + o(x^n)$; $\ln(1+x) = \sum_{k=1}^n (-1)^{k+1}\frac{x^k}{k} + o(x^n)$; $\sin(x) = \sum_{k=0}^n (-1)^k \frac{x^{2k+1}}{(2k+1)!} + o(x^{2n+1})$; $\cos(x) = \sum_{k=0}^n (-1)^k \frac{x^{2k}}{(2k)!} + o(x^{2n})$; $\frac{1}{1-x} = \sum_{k=0}^n x^k + o(x^n)$