6.7 to 8.4

Order statistics: $Y_{(1)} = \min\{Y_1, Y_2, \dots Y_n\}$; $Y_{(n)} = \max\{Y_1, Y_2, \dots Y_n\}$; $f_{Y_{(1)}}(y) = n\big(1 - F(y)\big)^{n-1} f(y)$;
$f_{Y_{(n)}}(y) = n\big(F(y)\big)^{n-1} f(y)$; $f_{Y_{(i)}}(y) = \dfrac{n!}{(i-1)!\,(n-i)!}\big(F(y)\big)^{i-1}\big(1-F(y)\big)^{n-i} f(y)$

---

random sample: iid. Statistic: real-valued function of observable rvs. Parameter: "truth about the world."
Statistic:sample::parameter:population. Sampling distribution: prob dist of statistic (Xbar).

$Y: rv, pdf: f_Y(y), h(Y): strictly\ incr\ or\ decr\ funct\ of\ Y; pdf\ of\ U = h(Y): f_U(u) = f_Y\big(h^{-1}(u)\big)\left|\frac{dy}{du}\right|$
$U = a_1Y_1 + a_2Y_2 + \cdots + a_nY_n, \Rightarrow\ m_U(t) = m_{Y_1}(ta_1) * m_{Y_1}(ta_1) + \cdots + m_{Y_n}(ta_n)$

$U = a_1X_1 + \cdots + a_nX_n \sim N(\mu_U, \sigma_U^2), where\ \mu_U = a_1\mu_1 + \cdots + a_n\mu_n,$ and $\sigma_U^2 = a_1^2\sigma_1^2 + \cdots + a_n^2\sigma_n^2$ Ex? 7p14

$U = Z_1^2 + \cdots + Z_n^2 \sim \chi_\nu^2 \sim Gamma(\alpha = \nu/2, \beta = 2)$

$Smpl\ mean: \bar{X} = \frac{1}{n}(X_1 + \cdots + X_n), Smpl\ var: S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 \therefore \frac{(n-1)S^2}{\sigma^2} = \frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{\sigma^2} \sim \chi_{n-1}^2$

$X \sim \chi_\nu^2: P(X(\geq \chi_{\nu,\alpha}^2) = \alpha$ $\quad \chi_{\nu,\alpha}^2$ is α^{th} percentage point and $(1-\alpha)^{th}$ quantile of χ^2 with ν deg. fr

---

$X_1, \dots X_n: i.i.d\ N(\mu, \sigma^2)\ rvs \Rightarrow \bar{X}, S^2: independent \quad \bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right), \frac{(n-1)S^2}{\sigma^2} \sim \chi_{n-1}^2$

$Z \sim N(0,1)\ and\ \chi_\nu^2 \sim chi.sq.with\ \nu\ deg.fr\ and\ Z, \chi_\nu^2: ind \Rightarrow t = t_\nu = \frac{Z}{\sqrt{((\chi_\nu^2)/\nu)}}\ \ pdf: f(t) = K\left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}},$

$-\infty < t < \infty\ \ where\ K = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)}$ Properties: mean=0, similar shape to N, but with fatter tails (larger variation) $\quad As\ \nu \to \infty, t_\nu \to N(0,1) \quad E(t_\nu) = 0\ [if\ \nu > 1] \quad V(t_\nu) = \frac{\nu}{\nu-2}\ \ [if\ \nu > 2]$
$t_1: Cauchy(0,1)\ dist \quad [no\ defined\ mean, var, or\ higher\ moments] \quad P\big(t_\nu \geq t_{\nu,\alpha}\big) = \alpha$
$t_{\nu,\alpha}$ is α^{th} percentage point and $(1-\alpha)^{th}$ quantile of t with ν deg. fr $\quad If\ \bar{X}: mean\ and\ S^2: samp\ var$ $of\ rand\ samp\ from\ N(\mu, \sigma^2), then\ \ Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)\ \ W = \frac{(n-1)S^2}{\sigma^2} \sim \chi_{n-1}^2$ and $T = \frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t_{n-1}$

$If\ W_{1,2}\ are\ ind.chi.sq\ rvs\ with\ \nu_{1,2}\ deg.fr, then\ F = \frac{W_1/\nu_1}{W_2/\nu_2}\ \ follows\ Fdist\ with\ \nu_1\ num.deg.fr\ and$ $\nu_2\ denom.deg.fr \quad F \sim F(\nu_1, \nu_2)\ \ pdf: f_F(x) = \frac{\Gamma\left(\frac{\nu_1+\nu_2}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)}\left(\frac{\nu_1}{\nu_2}\right)x^{\frac{\nu_1}{2}-1}\left(1 + \frac{\nu_1}{\nu_2}x\right)^{-\left(\frac{\nu_1+\nu_2}{2}\right)} \quad for\ \ 0 < x < \infty$
$\nu_1, \nu_2: \mathbb{Z}_{>0} \quad P\big(F \geq F_{\nu_1,\nu_2,\alpha}\big) = \alpha \quad For\ \nu_2 > 2: E\big(F_{\nu_1,\nu_2}\big) = \frac{\nu_2}{\nu_2-2} \quad F \sim F\big(F_{\nu_1,\nu_2}\big) \Rightarrow \frac{1}{F} \sim F\big(F_{\nu_2,\nu_1}\big)$
$F_{\nu_1,\nu_2,1-\alpha} = \frac{1}{F_{\nu_1,\nu_2,\alpha}} \quad If\ X_{11}, \dots X_{1n_1}\ ind.rand.samps\ from\ N(\mu_1, \sigma^2)\ and\ X_{21}, \dots X_{2n_2}\ ind.rand.samps$ $from\ N(\mu_2, \sigma^2), then (S_1^2)/(S_2^2) \sim F_{n_1-1,n_2-1}$

CLT $\quad If\ X_1, \dots, X_n: iid\ rvs\ with\ finite\ mean\ \mu\ and\ var\ \sigma^2 < \infty\ \ Then, as\ n \to \infty, the\ dist\ of\ U_n = \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$ $approaches\ that\ of\ N(0,1) \quad Thus, for\ large\ n, \quad P\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \leq z\right) \approx P(Z \leq z) \quad With\ CLT, \bar{X}\ is$ $asymptotically\ normal\ with\ mean\ \mu\ and\ variance\ \frac{\sigma^2}{n} \quad \left[written\ as\ \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \dot{\sim} N(0,1)\right]$

Continuity correction: if X is discrete, but approx. normal, add 0.5 to Xbar in calculation:
$X \dot{\sim} N(125,25) \to P(X \leq 130) = P(X < 131) \approx P\left(Z \leq \frac{130.5-125}{5}\right) \dots$ Binomial approximation:
$Y \sim Bin(n,p)\ \ by\ CLT: Y = n\bar{X} \dot{\sim} N\big(np, np(1-p)\big) \therefore if\ n: large, P(Y \leq y) \approx P\left(Z \leq \frac{(y+0.5)-np}{\sqrt{np(1-p)}}\right)$ [w/CC]
nec crit for BinApprx: $0 < p - 3\sqrt{p(1-p)/n}$ and $p + 3\sqrt{p(1-p)/n} < 1$ [n:large and p not near 0 or 1]

Estimator $\hat{\theta}$ of unknown parameter θ: statistic (function of data), rv with prob dist (sampling dist) and moments $E(\hat{\theta})$ *and* $V(\hat{\theta}) = E\left(\left(\hat{\theta} - E(\hat{\theta})\right)^2\right)$ and std dev (std err) $\sqrt{V(\hat{\theta})} = SD(\hat{\theta}) = SE(\hat{\theta})$

$Unbiased: E(\hat{\theta}) = \theta, for\ all\ values\ of\ \theta$ $\quad Bias(\hat{\theta}) = E(\hat{\theta}) - \theta$ $\quad Precision: measured\ with\ SE$

$Mean\ Squared\ Error: MSE(\hat{\theta}) = E\left((\hat{\theta} - \theta)^2\right) = V(\hat{\theta}) + \left(Bias(\hat{\theta})\right)^2$ $\quad RE(\hat{\theta}_1, \hat{\theta}_2) = \frac{MSE(\hat{\theta}_2)}{MSE(\hat{\theta}_1)}$

$absolute\ error\ of\ estimation: \varepsilon = |\hat{\theta} - \theta|$ $\quad Ex: P(\varepsilon \le b) = .9 \rightarrow P(-b \le \hat{\theta} - \theta \le b) = .9$

$Chebysheff's\ Theorem: P(|Y - E(Y)| < kSD(Y)) \ge 1 - \frac{1}{k^2}$

$Empirical\ Rule: If\ \hat{\theta} \sim N\ and\ Bias(\hat{\theta}) = 0, then\ P\left(\varepsilon < 2SE(\hat{\theta})\right) \approx .95\ and\ P\left(\varepsilon < 1.645SE(\hat{\theta})\right) \approx .90$