

# Topics in Statistical Theory — Example Sheet 1

Lucas Riedstra

...

**Question 1.** Let  $U_1, \dots, U_n \stackrel{\text{iid}}{\sim} U(0, 1)$  and let  $Y_1, \dots, Y_{n+1} \stackrel{\text{iid}}{\sim} \text{Exp}(1)$ . Writing  $S_j := \sum_{i=1}^j Y_i$  for  $j = 1, \dots, n+1$ , show that

$$U_{(j)} \stackrel{d}{=} \frac{S_j}{S_{n+1}} \sim \text{Beta}(j, n-j+1)$$

for  $j = 1, \dots, n$ .

*Solution.* We compute the density function of  $U_{(j)}$  as follows: let  $x \in (0, 1)$ , then we know that

$$f_{(j)}(x) = \frac{d}{dx} F_{(j)}(x) = \lim_{h \rightarrow 0} \frac{F_{(j)}(x+h) - F_{(j)}(x)}{h} = \lim_{h \rightarrow 0} \frac{\mathbb{P}(x < U_{(j)} \leq x+h)}{h}.$$

The probability  $\mathbb{P}(x < U_{(j)} \leq x+h)$  is the probability that exactly  $j-1$  of the  $U_i$  are less than  $x$ , and that at least one of the  $U_i$  is in  $(x, x+h]$ .

The probability that two or more of the  $U_i$  lie in  $(x, x+h]$  is  $O(h^2)$  and therefore negligible, so we must compute the probability that exactly  $j-1$  of the  $U_i$  are smaller than  $x$ , one of the  $U_i$  is in  $(x, x+h]$ , and the other  $U_i$  are greater than  $x+h$ . This is easily seen to be

$$\begin{aligned} & \binom{n}{j-1} \mathbb{P}(U \leq x)^{j-1} \cdot (n-j+1) \mathbb{P}(x < U \leq x+h) \cdot \mathbb{P}(U > x+h)^{n-j} \\ &= \frac{n!}{(j-1)!(n-j+1)!} (n-j+1) x^{j-1} h (1-x-h)^{n-j} \\ &= \frac{n!}{(j-1)!(n-j)!} x^{j-1} (1-x-h)^{n-j} h. \end{aligned}$$

Therefore, we easily compute

$$f_{(j)}(x) = \lim_{h \rightarrow 0} \frac{\frac{n!}{(j-1)!(n-j)!} x^{j-1} (1-x-h)^{n-j} h}{h} = \frac{n!}{(j-1)!(n-j)!} x^{j-1} (1-x)^{n-j}.$$

This is also the density function of a  $\text{Beta}(j, n-j+1)$  distribution.

Finally, define  $T_j = S_{n+1} - S_j$ , so that  $S_j$  and  $T_j$  are independent. It is known that  $S_j \sim \text{Gamma}(j, 1)$ ,  $T_j \sim \gamma(n-j+1, 1)$ , and furthermore that

$$\frac{S_j}{S_{n+1}} = \frac{S_j}{S_j + T} \stackrel{d}{=} \frac{\Gamma(j, 1)}{\Gamma(j, 1) + \Gamma(n-j+1, 1)} \sim \text{Beta}(j, n-j+1).$$

**Question 2.** Let  $X$  be a random variable with mean zero that satisfies  $a \leq X \leq b$ . Use convexity to show that for every  $t \in \mathbb{R}$ , we have

$$\log \mathbb{E}(e^{tX}) \leq -\alpha u + \log(\beta + \alpha e^u),$$

where  $u := t(b-a)$  and  $\alpha := 1 - \beta = -a/(b-a)$ . Using a second-order Taylor expansion around the origin, deduce that  $\log \mathbb{E}(e^{tX}) \leq t^2(b-a)^2/8$ .

*Proof.* Let  $x \in [a, b]$ , then we know there exists a unique  $\lambda \in [0, 1]$  such that  $x = (1 - \lambda)a + \lambda b$ . A simple computation gives  $\lambda = (x - a)/(b - a)$ ,  $1 - \lambda = (b - x)/(b - a)$ . By convexity of  $t \mapsto e^{tx}$  we find

$$e^{tx} \leq \frac{b - x}{b - a} e^{ta} + \frac{x - a}{b - a} e^{tb}.$$

From this we deduce that

$$\mathbb{E}[e^{tX}] \leq \mathbb{E}\left[\frac{b - X}{b - a} e^{ta} + \frac{X - a}{b - a} e^{tb}\right] = \frac{b}{b - a} e^{ta} + \frac{-a}{b - a} e^{tb} = \beta e^{ta} + \alpha e^{tb} = e^{-\alpha u + \log(\beta + \alpha e^u)}.$$

Since  $\log$  is increasing, we can take the logarithm on both sides to conclude

$$\log \mathbb{E}[e^{tX}] \leq -\alpha u + \log(\beta + \alpha e^u).$$

Now, we compute the Taylor polynomial of  $f(u) := -\alpha u + \log(\beta + \alpha e^u)$  in  $u = 0$ : we have

$$\begin{aligned} f(0) &= \log(\beta + \alpha) = \log(1) = 0; \\ f'(u) &= -\alpha + \frac{\alpha e^u}{\beta + \alpha e^u}; \\ f'(0) &= -\alpha + \frac{\alpha}{\beta + \alpha} = 0; \\ f''(u) &= \frac{(\beta + \alpha e^u)\alpha e^u - (\alpha e^u)^2}{(\beta + \alpha e^u)^2} = \frac{\alpha e^u}{(\beta + \alpha e^u)} \left(1 - \frac{\alpha e^u}{(\beta + \alpha e^u)}\right) \end{aligned}$$

Note that  $\frac{\alpha e^u}{\beta + \alpha e^u} \in [0, 1]$  since  $\alpha, \beta \geq 0$  (this holds because  $a$  must be negative and  $b$  must be positive due to the condition  $\mathbb{E}X = 0$ ). For  $y \in [0, 1]$ , the polynomial  $y(1 - y)$  takes values in  $[0, \frac{1}{4}]$ . Therefore, by Taylor's theorem with remainder, we conclude

$$\log \mathbb{E}[e^{tX}] \leq \left(\sup_{u \in \mathbb{R}} \frac{f''(u)}{2}\right) u^2 = \frac{1}{8} u^2 = \frac{t^2(b - a)^2}{8}.$$

□

**Question 3.** Let  $X_1, \dots, X_n$  be independent with distribution  $P$  on a measurable space  $(\mathcal{X}, \mathcal{A})$ , and let  $\hat{P}_n$  be the empirical measure of  $X_1, \dots, X_n$ ; thus  $\hat{P}_n(A) = n^{-1} \sum_{i=1}^n \mathbb{1}_{U_i \in A}$ . Show that, for all  $\varepsilon > 0$  and  $A \in \mathcal{A}$ , we have

$$\mathbb{P}\left(\left|\hat{P}_n(A) - P(A)\right| > \varepsilon\right) \leq 2e^{-2n\varepsilon^2}.$$

*Proof.* Define a new distribution  $Y = \mathbb{1}_{X \notin A}$ . Its distribution function is given by

$$F_Y(y) = \begin{cases} 0 & y < 0; \\ P(A) & y \in [0, 1); \\ 1 & y \geq 1. \end{cases}$$

The empirical distribution function of  $Y_1, \dots, Y_n \stackrel{\text{iid}}{\sim} Y$  is given by

$$\hat{F}_n(y) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{Y_i \leq y},$$

and thus for  $y \in [0, 1)$  we have

$$\hat{F}_n(y) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{Y_i \leq y} = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{X_i \in A} = \hat{P}_n(A).$$

By the DKW inequality we find

$$\mathbb{P}\left(\left|\hat{P}_n(A) - P(A)\right| > \varepsilon\right) = \mathbb{P}\left(\sup_{y \in \mathbb{R}} \left|\hat{F}_n(y) - F(y)\right| > \varepsilon\right) \leq 2e^{-2n\varepsilon^2}.$$

□

**Question 4.** Let  $X \sim \text{Bin}(n, p)$ . Compare the Hoeffding, Bennett, and Bernstein upper bounds on  $\mathbb{P}(X/n \geq \frac{1}{2})$  as  $p \rightarrow 0$ .

*Solution.* Note that  $X/n$  is the average of  $n$  i.i.d. random variables  $Y_i \sim \text{Bern}(p)$ , where  $Y_i \in [0, 1]$  for all  $i$ .

We start with Hoeffding's inequality. In this case, we have

$$\mathbb{P}(X/n - p \geq x) \leq \exp\left(-\frac{2n^2(1/2)^2}{n}\right) = \exp\left(-\frac{n}{2}\right),$$

and this bound also holds as  $p \rightarrow 0$ .

We continue with Bennett's inequality. We consider the random variables  $Y_i - p$ , which are bounded from above by  $b = 1 - p$ . Now Bennett's inequality tells us, with  $\sigma_p^2 = \text{Var}(Y_i - p) = p(1 - p)$  that

$$\mathbb{P}(X/n \geq x) \leq \exp\left(-\frac{n}{(1-p)^2} h\left(\frac{1-p}{2p(1-p)}\right)\right) = \exp\left(-\frac{n}{(1-p)^2} h\left(\frac{1}{2p}\right)\right).$$

We compute

$$h\left(\frac{1}{2p}\right) = \left(1 + \frac{1}{2p}\right) \log\left(1 + \frac{1}{2p}\right) - \frac{1}{2p} = \log\left(1 + \frac{1}{2p}\right) + \frac{1}{2p} \left(\log\left(1 + \frac{1}{2p}\right) - 1\right) \xrightarrow{p \downarrow 0} +\infty.$$

Since  $\frac{n}{(1-p)^2}$  is clearly bounded by  $n$ , we conclude that

$$\mathbb{P}(X/n \geq x) \rightarrow e^{-\infty} = 0.$$

We finish with Bernstein's inequality. We have for  $q \geq 3$  and  $p \leq \frac{1}{2}$  that

$$\begin{aligned} \mathbb{E}[(Y_i - p)_+^q] &= p(1-p)^q = \sigma_p^2(1-p)^{q-1} = (q!\sigma_p^2(1-p)^{q-2}/2) \cdot (2(1-p)/q!) \\ &\leq q!\sigma_p^2((1-p)/3)^{q-2}/2. \end{aligned}$$

Now Bernstein's inequality tells us that

$$\mathbb{P}(X/n \geq \frac{1}{2}) \leq \exp\left(-\frac{n(1/2)^2}{2(\sigma_p^2 + (1-p)/6)}\right) = \exp\left(-\frac{n}{8\sigma_p^2 + 4(1-p)/3}\right) \xrightarrow{p \rightarrow 0} \exp\left(-\frac{3n}{4}\right).$$

Of course, the true limit is 0 for any  $n$ , which is only given by Bennett's inequality. We also see that Hoeffding's inequality gives the worst result.

**Question 5.** Derive the following alternative form of Bernstein's inequality: under the same conditions,

$$\mathbb{P}\left(\bar{X} \geq \sqrt{\frac{2\sigma^2 \log(1/\delta)}{n}} + \frac{c}{n} \log(1/\delta)\right) \leq \delta$$

for every  $\delta \in (0, 1]$ .

*Proof.* We have

$$\mathbb{P}(\bar{X} \geq x) \leq \exp\left(-\frac{nx^2}{2(\sigma^2 + cx)}\right) =: \delta.$$

Now we just need express  $x$  in terms of  $\delta$ : taking logarithms on both sides we obtain

$$-\frac{nx^2}{2(\sigma^2 + cx)} = \log(\delta) \implies nx^2 = 2(\sigma^2 + cx) \log(1/\delta) \implies nx^2 - 2c \log(1/\delta)x - 2\sigma^2 \log(1/\delta) = 0.$$

Using the abc-formula with the fact that  $x \geq 0$  yields

$$\begin{aligned} x &= \frac{2c \log(1/\delta) + \sqrt{4c^2 \log^2(1/\delta) + 8n\sigma^2 \log(1/\delta)}}{2n} \\ &= \frac{c}{n} \log(1/\delta) + \sqrt{\frac{c^2}{n^2} \log^2(1/\delta) + \frac{2\sigma^2}{n} \log(1/\delta)} \\ &= \text{????} \end{aligned}$$

□

**Question 6.** (a) Let  $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} F$  and let  $\hat{F}_n$  denote their empirical distribution function. For  $t_1 < \dots < t_k$ , write down the distribution of

$$n(\hat{F}_n(t_1), \hat{F}_n(t_2) - \hat{F}_n(t_1), \dots, \hat{F}_n(t_k) - \hat{F}_n(t_{k-1}), 1 - \hat{F}_n(t_k)).$$

(b) Find the asymptotic distribution of  $n^{1/2}(\hat{F}_n(t_1) - F(t_1), \dots, \hat{F}_n(t_k) - F(t_k))$ .

*Solution.* (a) Write  $n\hat{F}_n(t) = \sum_{i=1}^n \mathbb{1}_{X_i \leq t} = \#\{i \mid X_i \leq t\}$ , and analogously, for  $t < u$ ,  $n(\hat{F}_n(u) - \hat{F}_n(t)) = \#\{i \mid t < X_i \leq u\}$ .

Then, defining  $t_0 = -\infty$  and  $t_{k+1} = \infty$ , we find that

$$\begin{aligned} \mathbb{P}\left[n(\hat{F}_n(t_1), \dots, 1 - \hat{F}_n(t_k)) = (a_1, \dots, a_{k+1})\right] \\ = \mathbb{P}[\text{exactly } a_i \text{ of the } X_i \text{ lie in } (t_{i-1}, t_i] \text{ for } i = 1, \dots, n]. \end{aligned}$$

In this case, we have a multinomial distribution with  $n$  trials and probabilities  $F(t_1), F(t_2) - F(t_1), \dots, F(t_k) - F(t_{k-1}), 1 - F(t_k)$ . Therefore, the probability is 0 if  $\sum_i a_i \neq n$  and else it is

$$\frac{n!}{a_1! \dots a_{k+1}!} F(t_1)^{a_1} \dots (1 - F(t_k))^{a_{k+1}}.$$

(b) By the central limit theorem, the asymptotic distribution is  $N(0, \Sigma)$ , where  $\Sigma$  is the covariance matrix of  $(\hat{F}_n(t_1), \dots, \hat{F}_n(t_k))$ . We will compute the entries of  $\Sigma$ .

Choose  $t \in \mathbb{R}$  arbitrarily. Then we have

$$\begin{aligned} \text{Var}(\hat{F}_n(t)) &= \mathbb{E}[\hat{F}_n^2(t)] - \mathbb{E}[\hat{F}_n(t)]^2 = \mathbb{E}\left[\left(\frac{1}{n} \sum_i \mathbb{1}_{X_i \leq t}\right)^2\right] - F^2(t) \\ &= \frac{1}{n^2} \mathbb{E}\left[\sum_i \mathbb{1}_{X_i \leq t} + 2 \sum_{i < j} \mathbb{1}_{X_i \leq t} \mathbb{1}_{X_j \leq t}\right] - F^2(t) \\ &= \frac{F(t) + (n-1)F^2(t)}{n} - F^2(t) = \frac{F(t)(1 - F(t))}{n}, \end{aligned}$$

so we have computed the diagonal entries  $\Sigma_{ii} = \frac{F(t_i)(1-F(t_i))}{n}$ .

Now we must compute the covariances: assume  $s < t$ , then

$$\begin{aligned} \text{Cov}(\hat{F}_n(s), \hat{F}_n(t)) &= \mathbb{E}[\hat{F}_n(s)\hat{F}_n(t)] - \mathbb{E}[\hat{F}_n(s)]\mathbb{E}[\hat{F}_n(t)] \\ &= \frac{1}{n^2} \sum_{i,j} \mathbb{E}[\mathbb{1}_{X_i \leq s} \mathbb{1}_{X_j \leq t}] - F(s)F(t) \\ &= \frac{1}{n^2} (nF(s) + n(n-1)F(s)F(t)) - F(s)F(t) \\ &= \frac{F(s) + (n-1)F(s)F(t)}{n} - F(s)F(t) = \frac{F(s) - F(s)F(t)}{n}. \end{aligned}$$

This gives the diagonal entries  $\Sigma_{ij} = \frac{F(t_i) - F(t_i)F(t_j)}{n}$  for  $i < j$ . In the end, we find

$$\Sigma_{ij} = \frac{1}{n} \cdot \begin{cases} F(t_i)(1 - F(t_i)) & \text{if } i = j, \\ F(t_{\min(i,j)}) - F(t_i)F(t_j) & \text{if } i \neq j. \end{cases}$$

**Question 7.** We say that a continuous process  $(B_t)_{t \in [0,1]}$  is a standard Brownian motion on  $[0,1]$  if  $B_0 = 0$  and if, for  $0 \leq s_1 \leq t_1 \leq \dots \leq s_k \leq t_k \leq 1$ , we have  $(B_{t_1} - B_{s_1}, \dots, B_{t_k} - B_{s_k}) \sim N_k(0, \Sigma)$ , where  $\Sigma := \text{diag}(t_1 - s_1, \dots, t_k - s_k)$ . The process  $(W_t)_{t \in [0,1]}$  defined by  $W_t := B_t - tB_1$  is called a Brownian bridge, or tied-down Brownian motion, because  $W_0 = W_1 = 0$ . Compute the distribution of  $(W_{t_1}, \dots, W_{t_k})$ .

*Solution.* Note that  $W_t = B_t - tB_1 = (1-t)(B_t - B_0) - t(B_1 - B_t)$ . Now, since  $(B_t - B_0)$  and  $(B_1 - B_t)$  are independent with distributions  $N(0, t)$  and  $N(0, 1-t)$  distributions respectively, we find that

$$W_t \sim (1-t)N(0, t) + tN(0, 1-t) = N\left(0, \frac{t}{\sqrt{1-t}}\right) + N\left(0, \frac{1-t}{\sqrt{t}}\right) = N\left(0, \frac{t}{\sqrt{1-t}} + \frac{1-t}{\sqrt{t}}\right).$$

??

**Question 8.** Let  $\varphi$  denote the standard normal density function, which is a bounded, second-order kernel. For  $r \in \mathbb{N}_0$ , define the  $r$ -th Hermite polynomial  $H_r$  by  $H_r(x) := (-1)^r \varphi^{(r)}(x) / \varphi(x)$ . Prove that  $H_r$  is a monic polynomial of degree  $r$  that is even if  $r$  is even and odd if  $r$  is odd. Show further that

$$\int_{-\infty}^{\infty} H_r(u) H_s(u) \varphi(u) du = \begin{cases} (2\pi)^{1/2} r!, & r = s, \\ 0, & r \neq s. \end{cases}$$

Now fix an integer  $\ell \geq 2$  and define

$$K_\ell(u) := \sum_{r=0}^{\ell-1} \frac{H_r(0) H_r(u)}{(2\pi)^{1/2} r!} e^{-u^2/2}.$$

Prove that  $K_\ell$  is a bounded kernel of order  $\ell$ .

*Proof.* We prove this by induction on  $r$ . For  $r = 0$ , we have  $H_0(x) = 1$ , which is indeed an even monic polynomial of degree 0. Now, suppose the claim holds for a given  $r$ , that is,  $H_r(x) = (-1)^r \varphi^{(r)}(x) / \varphi(x) = p(x)$  for some monic polynomial  $p$  of degree  $r$ , which is even if  $r$  is even and odd if  $r$  is odd. Then we have

$$\begin{aligned} \varphi^{(r)}(x) &= (-1)^r p(x) \varphi(x) = (-1)^r (2\pi)^{-1/2} p(x) \exp(-x^2/2) \\ \varphi^{(r+1)}(x) &= (-1)^{r+1} 2\pi^{-1/2} (p'(x) - xp(x)) \exp(-x^2/2) \\ H_{r+1}(x) &= (-1)^{r+1} (-1)^{r+1} (p'(x) - xp(x)) = xp(x) - p'(x). \end{aligned}$$

Now it is clear that  $H$  is a monic polynomial of degree  $r$  since  $p$  was assumed monic. Furthermore, since derivatives of even functions are odd and vice versa, it is clear that  $H$  is odd if  $p$  is even and vice versa.

Now, suppose  $r < s$ , then

$$\int_{-\infty}^{\infty} H_r(u) H_s(u) \varphi(u) \, du = (-1)^s \int_{-\infty}^{\infty} H_r(u) \varphi^{(s)}(u) \, du \stackrel{\text{IBP}}{=} \int_{-\infty}^{\infty} H_r^{(s)}(u) \varphi(u) \, du = 0,$$

since  $H_r^{(s)} = 0$  if  $r < s$ .

However, if  $r = s$ , then following the same line of reasoning as above and using the fact that  $H_r^{(r)} = r!$ , we find

$$\int_{-\infty}^{\infty} H_r^2(u) \varphi(u) \, du = r! \int_{-\infty}^{\infty} \varphi(u) \, du = r!.$$

Now we consider  $K_\ell$ : we have

$$\int_{-\infty}^{\infty} K_\ell(u) \, du = \sum_{r=0}^{\ell-1} \frac{H_r(0)}{r!} \int_{-\infty}^{\infty} H_r(u) \varphi(u) \, du = \sum_{r=0}^{\ell-1} \frac{(-1)^r H_r(0)}{r!} \int_{-\infty}^{\infty} \varphi^{(r)}(u) \, du.$$

Note that every term in the above sum vanishes except for the  $r = 0$  term due to the integral, and the  $r = 0$  term is 1, so  $K_\ell$  is indeed a kernel.

We verify that  $K_\ell$  has order  $\ell$ : let  $j \in \{1, \dots, \ell - 1\}$ , then we have

$$\int_{-\infty}^{\infty} u^j K_\ell(u) \, du = \sum_{r=0}^{\ell-1} \frac{(-1)^r H_r(0)}{r!} \int_{-\infty}^{\infty} u^j \varphi^{(r)}(u) \, du.$$

Clearly, if  $j < r$ , integral in the sum vanishes. If  $j = r$ , the integral yields  $r!$ . So we have

$$\int_{-\infty}^{\infty} u^j K_\ell(u) \, du = (-1)^j H_j(0) + \sum_{r=0}^{j-1} \int_{-\infty}^{\infty} u^j \varphi^{(r)}(u) \, du.$$

□