

# UIT-T

SECTEUR DE LA NORMALISATION  
DES TÉLÉCOMMUNICATIONS  
DE L'UIT

# E.840

(06/2018)

## SÉRIE E: EXPLOITATION GÉNÉRALE DU RÉSEAU, SERVICE TÉLÉPHONIQUE, EXPLOITATION DES SERVICES ET FACTEURS HUMAINS

Qualité des services de télécommunication: concepts,  
modèles, objectifs et planification de la sûreté de  
fonctionnement – Modèles pour les services de  
télécommunication

---

### **Cadre statistique applicable à la notation et au classement comparatifs de la qualité de fonctionnement de réseau de bout en bout**

Recommandation UIT-T E.840

RECOMMANDATIONS UIT-T DE LA SÉRIE E  
**EXPLOITATION GÉNÉRALE DU RÉSEAU, SERVICE TÉLÉPHONIQUE, EXPLOITATION DES  
SERVICES ET FACTEURS HUMAINS**

<b>EXPLOITATION DES RELATIONS INTERNATIONALES</b>	
Définitions	E.100–E.103
Dispositions de caractère général concernant les Administrations	E.104–E.119
Dispositions de caractère général concernant les usagers	E.120–E.139
Exploitation des relations téléphoniques internationales	E.140–E.159
Plan de numérotage du service téléphonique international	E.160–E.169
Plan d'acheminement international	E.170–E.179
Tonalités utilisées dans les systèmes nationaux de signalisation	E.180–E.189
Plan de numérotage du service téléphonique international	E.190–E.199
Service mobile maritime et service mobile terrestre public	E.200–E.229
<b>DISPOSITIONS OPÉRATIONNELLES RELATIVES À LA TAXATION ET À LA  COMPTABILITÉ DANS LE SERVICE TÉLÉPHONIQUE INTERNATIONAL</b>	
Taxation dans les relations téléphoniques internationales	E.230–E.249
Mesure et enregistrement des durées de conversation aux fins de la comptabilité	E.260–E.269
<b>UTILISATION DU RÉSEAU TÉLÉPHONIQUE INTERNATIONAL POUR LES  APPLICATIONS NON TÉLÉPHONIQUES</b>	
Généralités	E.300–E.319
Phototélégraphie	E.320–E.329
<b>DISPOSITIONS DU RNIS CONCERNANT LES USAGERS</b>	E.330–E.349
<b>PLAN D'ACHEMINEMENT INTERNATIONAL</b>	E.350–E.399
<b>GESTION DE RÉSEAU</b>	
Statistiques relatives au service international	E.400–E.404
Gestion du réseau international	E.405–E.419
Contrôle de la qualité du service téléphonique international	E.420–E.489
<b>INGÉNIERIE DU TRAFIC</b>	
Mesure et enregistrement du trafic	E.490–E.505
Prévision du trafic	E.506–E.509
Détermination du nombre de circuits en exploitation manuelle	E.510–E.519
Détermination du nombre de circuits en exploitation automatique et semi-automatique	E.520–E.539
Niveau de service	E.540–E.599
Définitions	E.600–E.649
Ingénierie du trafic des réseaux à protocole Internet	E.650–E.699
Ingénierie du trafic RNIS	E.700–E.749
Ingénierie du trafic des réseaux mobiles	E.750–E.799
<b>QUALITÉ DES SERVICES DE TÉLÉCOMMUNICATION: CONCEPTS, MODÈLES,  OBJECTIFS ET PLANIFICATION DE LA SÛRETÉ DE FONCTIONNEMENT</b>	
Termes et définitions relatifs à la qualité des services de télécommunication	E.800–E.809
<b>Modèles pour les services de télécommunication</b>	<b>E.810–E.844</b>
Objectifs et concepts de qualité des services de télécommunication	E.845–E.859
Utilisation des objectifs de qualité de service pour la planification des réseaux de télécommunication	E.860–E.879
Collecte et évaluation de données d'exploitation sur la qualité des équipements, des réseaux et des services	E.880–E.899
<b>AUTRES</b>	E.900–E.999
<b>EXPLOITATION DES RELATIONS INTERNATIONALES</b>	
Plan de numérotage du service téléphonique international	E.1100–E.1199
<b>GESTION DES RÉSEAUX</b>	
Gestion des réseaux internationaux	E.4100–E.4199

*Pour plus de détails, voir la Liste des Recommandations de l'UIT-T.*

## Recommandation UIT-T E.840

### Cadre statistique applicable à la notation et au classement comparatifs de la qualité de fonctionnement de réseau de bout en bout

#### Résumé

La présente Recommandation est la première d'une série de Recommandations consacrée à l'évaluation comparative de la qualité de fonctionnement des réseaux de bout en bout. Elle présente un cadre d'analyse statistique pour l'évaluation comparative de la performance des réseaux et des services. Le cadre décrit des scénarios, des cas d'utilisation et des procédures d'évaluation comparative ainsi que des techniques statistiques pour classer les indicateurs fondamentaux de performance de qualité de bout en bout. La Recommandation fait référence aux services mobiles et aux campagnes d'évaluation comparative effectuées au moyen d'agents mobiles (dispositifs) dans des essais en voiture ou à pied, ainsi qu'au moyen d'agents fixes, à savoir des dispositifs installés dans des emplacements fixes (par exemple dans des centres commerciaux, des immeubles de bureaux ou des stades).

#### Historique

Édition	Recommandation	Approbation	Commission d'études	ID unique*
1.0	UIT-T E.840	13-06-2018	12	<a href="http://handle.itu.int/11.1002/1000/13621">11.1002/1000/13621</a>

#### Mots clés

Qualité de fonctionnement de bout en bout, évaluation comparative et classement de la qualité de fonctionnement des réseaux, cadre statistique.

---

\* Pour accéder à la Recommandation, reporter cet URL <http://handle.itu.int/> dans votre navigateur Web, suivi de l'identifiant unique, par exemple <http://handle.itu.int/11.1002/1000/11830-en>.

## AVANT-PROPOS

L'Union internationale des télécommunications (UIT) est une institution spécialisée des Nations Unies dans le domaine des télécommunications et des technologies de l'information et de la communication (ICT). Le Secteur de la normalisation des télécommunications (UIT-T) est un organe permanent de l'UIT. Il est chargé de l'étude des questions techniques, d'exploitation et de tarification, et émet à ce sujet des Recommandations en vue de la normalisation des télécommunications à l'échelle mondiale.

L'Assemblée mondiale de normalisation des télécommunications (AMNT), qui se réunit tous les quatre ans, détermine les thèmes d'étude à traiter par les Commissions d'études de l'UIT-T, lesquelles élaborent en retour des Recommandations sur ces thèmes.

L'approbation des Recommandations par les Membres de l'UIT-T s'effectue selon la procédure définie dans la Résolution 1 de l'AMNT.

Dans certains secteurs des technologies de l'information qui correspondent à la sphère de compétence de l'UIT-T, les normes nécessaires se préparent en collaboration avec l'ISO et la CEI.

## NOTE

Dans la présente Recommandation, l'expression "Administration" est utilisée pour désigner de façon abrégée aussi bien une administration de télécommunications qu'une exploitation reconnue.

Le respect de cette Recommandation se fait à titre volontaire. Cependant, il se peut que la Recommandation contienne certaines dispositions obligatoires (pour assurer, par exemple, l'interopérabilité et l'applicabilité) et on considère que la Recommandation est respectée lorsque toutes ces dispositions sont observées. Le futur d'obligation et les autres moyens d'expression de l'obligation comme le verbe "devoir" ainsi que leurs formes négatives servent à énoncer des prescriptions. L'utilisation de ces formes ne signifie pas qu'il est obligatoire de respecter la Recommandation.

## DROITS DE PROPRIÉTÉ INTELLECTUELLE

L'UIT attire l'attention sur la possibilité que l'application ou la mise en œuvre de la présente Recommandation puisse donner lieu à l'utilisation d'un droit de propriété intellectuelle. L'UIT ne prend pas position en ce qui concerne l'existence, la validité ou l'applicabilité des droits de propriété intellectuelle, qu'ils soient revendiqués par un membre de l'UIT ou par une tierce partie étrangère à la procédure d'élaboration des Recommandations.

A la date d'approbation de la présente Recommandation, l'UIT n'avait pas été avisée de l'existence d'une propriété intellectuelle protégée par des brevets à acquérir pour mettre en œuvre la présente Recommandation. Toutefois, comme il ne s'agit peut-être pas de renseignements les plus récents, il est vivement recommandé aux développeurs de consulter la base de données des brevets du TSB sous <http://www.itu.int/ITU-T/ipr/>.

© UIT 2020

Tous droits réservés. Aucune partie de cette publication ne peut être reproduite, par quelque procédé que ce soit, sans l'accord écrit préalable de l'UIT.

## TABLE DES MATIÈRES

	<b>Page</b>
1	Domaine d'application ..... 1
2	Références..... 1
3	Définitions ..... 1
4	Abréviations et acronymes ..... 1
5	Conventions ..... 2
6	Scénarios d'évaluation comparative..... 2
7	Conditions de l'évaluation comparative..... 3
8	Services faisant l'objet de l'évaluation comparative ..... 4
9	Cadre statistique..... 5
9.1	Nettoyage des données ..... 5
9.2	Distribution statistique des mesures ..... 5
9.3	Indicateurs statistiques de qualité de fonctionnement, erreurs types et signification statistique des résultats de l'évaluation comparative ..... 5
9.4	Notation et classement des KPI ou des KQI ..... 7
Annexe A	– Signification statistique à appliquer dans l'analyse de l'évaluation comparative des réseaux mobiles ..... 9
Annexe B	– Notation et classement statistiques de la qualité de fonctionnement d'un réseau 10
Appendice I	– Technique possible de notation et de classement statistiques d'un réseau ..... 12
Bibliographie	..... 14



## Recommandation UIT-T E.840

### Cadre statistique applicable à la notation et au classement comparatifs de la qualité de fonctionnement de réseau de bout en bout

#### 1 Domaine d'application

La présente Recommandation définit un cadre statistique, des scénarios d'évaluation comparative et des conditions d'application, dont les opérateurs et les régulateurs exigent l'utilisation lorsque la qualification et la quantification des différences de qualité de fonctionnement entre les indicateurs fondamentaux de performance (KPI) ou les indicateurs fondamentaux de qualité (KQI) de bout en bout affectent l'expérience des utilisateurs.

Cette Recommandation est nécessaire parce que, dans la course intense à la satisfaction d'utilisateurs existants de plus en plus exigeants et à l'élargissement simultané de leur base de clientèle à un coût optimal, les opérateurs ont amélioré le fonctionnement des réseaux dans une telle mesure que les différences entre eux deviennent de moins en moins marquées.

#### 2 Références

La présente Recommandation se réfère à certaines dispositions des Recommandations UIT-T et textes suivants qui, de ce fait, en sont partie intégrante. Les versions indiquées étaient en vigueur au moment de la publication de la présente Recommandation. Toute Recommandation ou tout texte étant sujet à révision, les utilisateurs de la présente Recommandation sont invités à se reporter, si possible, aux versions les plus récentes des références normatives suivantes. La liste des Recommandations de l'UIT-T en vigueur est régulièrement publiée. La référence à un document figurant dans la présente Recommandation ne donne pas à ce document, en tant que tel, le statut d'une Recommandation.

[UIT-T E.800]                      Recommandation UIT-T E.800 (2008), *Définitions des termes relatifs à la qualité des services*.

[UIT-T E.804]                      Recommandation UIT-T E.804 (2014), *Aspects de la qualité de service pour les services les plus prisés sur les réseaux mobiles*.

#### 3 Définitions

Aucune.

#### 4 Abréviations et acronymes

La présente Recommandation utilise les abréviations et acronymes suivants:

KPI     indicateur fondamental de performance (*key performance indicator*)

KQI     indicateur fondamental de qualité (*key quality indicator*)

MOS    note moyenne d'opinion (*mean opinion score*)

QoE    qualité d'expérience (*quality of experience*)

RF      fréquence radioélectrique (*radio frequency*)

TCP    protocole de commande de transmission (*transmission control protocol*)

## 5 Conventions

**5.2.1 StatScore:** note statistique, c'est-à-dire qualité globale de divers réseaux ou opérateurs par rapport au réseau présentant la meilleure qualité de fonctionnement. Le StatScore est calculé par service.

**5.2.2 GlobalNetScore:** note globale du réseau c'est-à-dire qualité globale de divers réseaux ou opérateurs par rapport au réseau présentant la meilleure qualité de fonctionnement. Le GlobalNetScore est calculé pour l'ensemble des services.

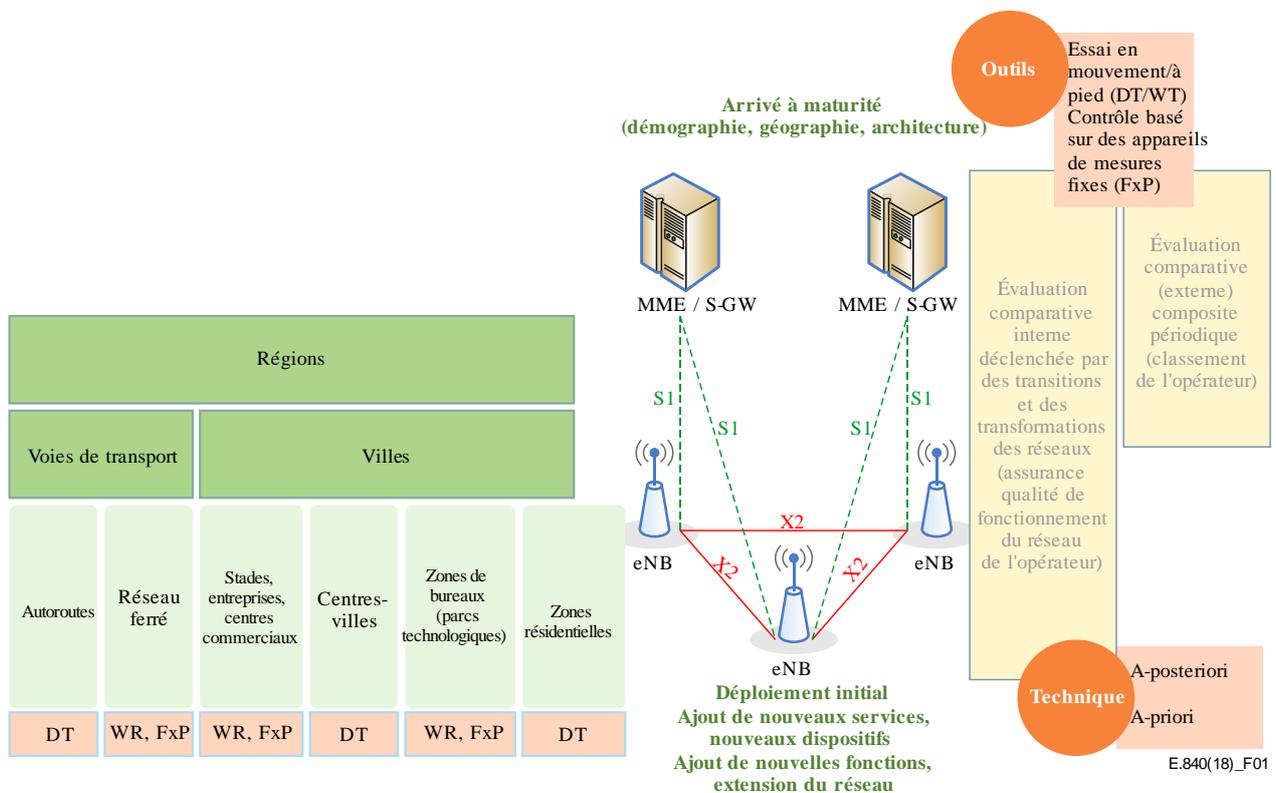
**5.2.3 StatDiff:** différence statistiquement significative entre deux indicateurs fondamentaux de performance (KPI) ou de qualité (KQI).

**5.2.4 THrelv:** différence minimale entre la valeur de deux indicateurs fondamentaux de performance (KPI) ou de qualité (KQI) pertinente pour l'utilisateur d'un service et au-delà de laquelle la signification statistique prévaut.

## 6 Scénarios d'évaluation comparative

En règle générale, l'évaluation comparative des réseaux présente deux cas d'utilisation principaux: interne et concurrentiel. L'évaluation comparative interne est axée sur l'assurance de continuité et de rentabilité de la qualité de fonctionnement d'un réseau et sur les améliorations nécessitant une évaluation à la fois lors du déploiement initial d'un réseau et pendant son développement, ainsi qu'à l'occasion du lancement de nouveaux services ou de nouveaux dispositifs. L'évaluation comparative interne s'effectue également dans les réseaux bien établis et parvenus à maturité. En outre, les régions dotées d'autoroutes, de villes et de lieux d'intérêt tels que des locaux professionnels, centres commerciaux, stades et bâtiments à usage résidentiel sont à prendre en compte dans les campagnes d'évaluation. L'évaluation comparative effectuée par les opérateurs eux-mêmes (ou par des prestataires de services pour leur compte) et par les régulateurs dans le but de vérifier l'état de la concurrence et de déterminer eux-mêmes leur classement couvre en général des régions, des voies de transport (réseaux routier et ferré) et des villes, et même des pays dans le cas de groupes d'opérateurs multinationaux, ainsi que des réseaux parvenus à maturité.

La Figure 1 présente la synthèse de ces cas d'utilisation ainsi que des types d'outils et de techniques recommandés. Dans le cas de zones situées en général en intérieur, telles que centres commerciaux, stades et locaux professionnels, l'évaluation comparative recourt souvent à des essais à pied. En dehors d'approches classiques telles que les campagnes d'essai en mouvement ou à pied, l'évaluation comparative interne et, dans une certaine mesure, concurrentielle, ainsi que les scénarios en intérieur ont intérêt à faire appel à des appareils de mesure fixes. Ceux-ci présentent en effet l'avantage d'une évolutivité rapide et à distance et de l'indépendance des dispositifs. Par conséquent, ces outils conviennent parfaitement aux essais en intérieur et aux nouveaux services lancés dans des lieux d'intérêt ainsi que dans des villes dans une certaine mesure. En outre, il est parfois possible d'appliquer des techniques d'analyse *a posteriori* ou *a priori*. Dans le premier cas, surtout utilisé dans l'évaluation comparative, la signification statistique des données collectées permet d'évaluer et de classer les notes des KPI ou des KQI de bout en bout. Par défaut, la précision de la mesure est intégrée au niveau de signification statistique. Les techniques *a priori* nécessitent le calcul préalable du nombre d'appareils de mesure requis pour satisfaire à une signification statistique et une précision de mesure spécifiées. Cette technique s'utilise en général lorsque les appareils de mesure sont onéreux ou lorsque la période de mesure est limitée.



**Figure 1 – Cas d'utilisation de l'évaluation comparative, types d'outils et de techniques recommandés**

## 7 Conditions de l'évaluation comparative

Quel que soit le cas d'utilisation, le cadre de l'évaluation comparative doit s'appuyer sur un ensemble de conditions préalables visant à garantir la cohérence, la validité, la fiabilité et la répétabilité. Le Tableau 1 décrit les prérequis de chaque phase de l'évaluation comparative: montage des équipements, configuration des essais, collecte, traitement et analyse des données. À noter qu'il s'agit des prérequis minimaux pour garantir la mise en place d'un environnement d'essai entièrement contrôlé ainsi que la validité de l'analyse statistique. Les détails spécifiques relatifs aux mesures figurent dans d'autres Recommandations UIT-T, telles que [UIT-T E.804].

**Tableau 1 – Prérequis minimaux**

Évaluation comparative	Prérequis de l'évaluation comparative
Montage des équipements	Le montage des équipements doit être homogène sur l'ensemble des réseaux, plateformes et dispositifs. Il faut utiliser les mêmes modèles d'appareils pour procéder à l'évaluation <i>comparative concurrentielle</i> ("comparaison à l'identique").
	Les conditions de fonctionnement des équipements doivent être conformes à celles décrites par leur fournisseur (par exemple, évitement de la surchauffe susceptible d'influer négativement sur le bon fonctionnement de l'appareil).

**Tableau 1 – Prérequis minimaux**

<b>Évaluation comparative</b>	<b>Prérequis de l'évaluation comparative</b>
Mise en place et configuration des essais	La mise en place de l'essai doit refléter l'expérience réelle de l'utilisateur. Des indications à ce sujet figurent dans [b-ETSI TR 102 581].
	Les paramètres de mise en place de l'essai doivent éviter les perturbations susceptibles d'exercer un impact négatif sur la qualité de fonctionnement du réseau: par exemple, il convient de vérifier l'emplacement du serveur de données et sa configuration avec les paramètres de configuration du protocole TCP afin d'assurer un bon débit à tous les opérateurs comparés. L'essai doit avoir lieu avant la collecte des données.
	La mise en place de l'essai doit prévoir plusieurs scénarios correspondant aux comportements des utilisateurs (par exemple, utilisation du protocole TCP, diverses tailles de fichiers, diverses longueurs d'appel vidéo ou voix) tout en en limitant le caractère intrusif afin de ne pas surcharger artificiellement le réseau.
Collecte des données	La collecte des données de mesure doit refléter l'expérience des utilisateurs (par exemple, note moyenne d'opinion (MOS) par service ainsi que les principaux KPI ou KQI de bout en bout influant sur lui). Les mesures doivent reposer sur des indicateurs ou des techniques de mesure adéquates conformes aux consignes des fournisseurs ou des Recommandations pertinentes.
	La collecte des données de mesure doit avoir lieu dans diverses conditions géographiques ou démographiques à différents moments (heure de pointe ou heure creuse, weekend ou jour de semaine, période de congés ou non).
Traitement et analyse des données	Comparer les données collectées avec des appareils identiques dans la même zone et pendant la même période (comparaison à l'identique).
	Utiliser la signification statistique pour effectuer des comparaisons utiles.
	Effectuer l'analyse par KPI ou KQI.

## 8 Services faisant l'objet de l'évaluation comparative

La liste des services mobiles traditionnellement faisant l'objet de l'évaluation comparative ainsi que leurs KPI et KQI et leurs points de déclenchement ne relèvent pas du champ d'application de la présente Recommandation. Voir [UIT-T E.800] et [UIT-T E.804] pour obtenir des détails complémentaires.

Si le but de l'évaluation de réseaux mobiles est de procéder à une analyse comparative détaillée par service, une tâche relevant en général de l'évaluation comparative interne (par exemple, lors de l'ajout d'un nouveau dispositif ou d'une nouvelle technologie), il est recommandé d'utiliser pour l'analyse un jeu complet de KPI et KQI de bout en bout (dans ce cas, les KPI sont les mesures obtenues à l'aide de modèles d'estimation de la qualité tels que [b-UIT-T P.863] pour la voix ou [b-UIT-T P.1203] pour le streaming vidéo). Il est également recommandé d'analyser les principales causes premières de l'éventuelle médiocrité de la qualité de fonctionnement à l'aide ce jeu d'indicateurs.

En revanche, si le but de l'évaluation comparative de réseaux mobiles est le classement de performance des KPI ou des KQI de bout en bout, une tâche relevant en général de l'évaluation comparative concurrentielle mais aussi de certains scénarios d'évaluation comparative interne (comparaison de marchés, évaluation périodique des performances du marché, par exemple), la notation et le classement peuvent faire appel à un jeu restreint de KPI ou de KQI influant sur la qualité d'expérience (QoE, voir [b-UIT-T P.10/G.100]) par service et pour l'ensemble des services faisant l'objet de l'évaluation comparative.

La présente Recommandation concerne l'évaluation comparative concurrentielle. D'autres Recommandations de l'UIT-T de la série consacrée à l'évaluation comparative abordent en détail ces ensembles de KPI ou de KQI.

## **9 Cadre statistique**

Le cadre recommandé vise à noter et à classer la qualité de fonctionnement de bout en bout des réseaux du point de vue de l'utilisateur et peut s'utiliser dans l'évaluation comparative interne et concurrentielle. Il définit les procédures de validation des données, des indicateurs d'évaluation statistiques et des tests de significativité ainsi que des consignes générales concernant la notation et de classement.

### **9.1 Nettoyage des données**

Afin que les résultats de l'évaluation comparative soient utiles et exacts, il convient de valider les données d'entrée de l'analyse. La validation consiste principalement à nettoyer les données, c'est-à-dire à supprimer les perturbations spécifiques aux mesures et les données incomplètes. Il est conseillé de remplacer les données manquantes par de nouvelles données collectées. S'il manque plus de 5% des données ou si ce pourcentage inclut des perturbations, il est recommandé de collecter de nouvelles données. Les perturbations propres aux mesures comprennent, entre autres: les appels silencieux dans le cas de services conversationnels audios ou vidéos, périodes d'indisponibilité des serveurs de données ou vidéo (par exemple, panne de serveur) dans le cas des services de streaming vidéo, valeurs MOS systématiquement très faibles inattendues pour la voix, la conversation vidéo ou le streaming. Dans ce cas, tout type de dégradation non causée par le réseau lui-même mais soit par les équipements de l'essai, soit par l'appareil de mesure doit être considéré comme une perturbation propre à la mesure.

Il convient d'éliminer ces perturbations des données de l'évaluation comparative afin de garantir la validité des hypothèses fournies par l'application des modèles statistiques.

### **9.2 Distribution statistique des mesures**

En général, la distribution statistique des valeurs d'un quelconque KPI ou KQI mesuré peut être estimée par une distribution gaussienne fondée sur le théorème central limite [b-Shaum]. En général, plus le nombre d'échantillons est élevé, plus la précision de l'estimation de la distribution gaussienne est bonne. Il est recommandé de vérifier les distributions statistiques de toutes les valeurs des KPI et des KQI. Deux méthodes sont possibles. L'une consiste à générer des graphiques de distribution des mesures du KPI ou du KQI analysé et à en vérifier la normalité par l'observation. L'autre consiste à utiliser des tests de validation pour vérifier la normalité tels que ceux de Kolmogorov-Smirnov, Anderson-Darling ou Shapiro-Wilk [b-Mehta]. En outre, dans le cas rare ou extrême de distributions non gaussiennes, il est possible de procéder à des essais non paramétriques. Les essais montrant les distributions expérimentales de divers KPI ou KQI, les essais de vérification de la normalité de la distribution ainsi que des cas spéciaux de distributions non gaussiennes sont abordés dans d'autres Recommandations de l'UIT-T de la série consacrée à l'évaluation comparative.

### **9.3 Indicateurs statistiques de qualité de fonctionnement, erreurs types et signification statistique des résultats de l'évaluation comparative**

#### **9.3.1 Indicateurs statistiques de qualité de fonctionnement**

L'analyse comparative doit reposer sur des indicateurs statistiques de qualité de fonctionnement reflétant la qualité de fonctionnement moyenne du réseau (représentée par des valeurs médianes,  $m$ ) ou sa cohérence (représentée par la probabilité que  $P_{th}$  soit supérieur à une valeur seuil prédéfinie). La présente Recommandation prend pour exemple l'indicateur statistique médian de qualité de fonctionnement. Des techniques similaires peuvent s'appliquer aux fins de cohérence.

### 9.3.2 Erreur type

Le calcul de l'erreur type à un niveau de confiance de 95% pour la moyenne ou Pth suppose une distribution gaussienne des KPI ou des KQI mesurés (voir le § 9.2).

Par conséquent, en fonction du type de KPI ou de KQI, de notes continues telles que les paramètres des fréquences radioélectriques (RF), des MOS ou de valeurs discrètes telles que le ratio succès/échec (r), l'erreur type à un niveau de confiance de 95% est donnée par les équations:

$$\text{StdError}(m) = z_{95\%} * \text{std} / \sqrt{N} = 1,96 * \text{std} / \sqrt{N}$$

$$\text{StdError}(r) = z_{95\%} * \sqrt{r * (1-r) / N} = 1,96 * \sqrt{r * (1-r) / N}$$

S'il y a moins de 30 échantillons, le quantile gaussien  $z_{95\%}$  doit être remplacé par la valeur tabulée Student  $t_{95\%}(N-1)$  où N représente le nombre d'échantillons disponibles.

NOTE – les erreurs types rendent compte de la précision de la mesure. Par conséquent, s'il faut obtenir une précision spécifique et si l'estimation de l'écart type est connue, le nombre minimum d'échantillons requis pour satisfaire cette précision au niveau de confiance sélectionné peut être calculé sur la base des équations du § 2. Cette méthode peut s'utiliser dans le cas d'une technique *a priori* telle qu'illustrée par la Figure 1 mais aussi dans le cadre de [b-UIT-T E.802] pour le calcul du nombre minimum d'échantillons.

### 9.3.3 Signification statistique

Les intervalles de confiance fournissent les marges d'erreur type à un niveau de confiance statistique donné, en général 95%. Cependant, deux KPI ou KQI comparés de valeurs proches et dont les intervalles de confiance se recoupent ne sont pas nécessairement statistiquement identiques. L'exactitude de la comparaison entre KPI ou KQI doit reposer sur la signification statistique afin de maintenir à 5% l'éventualité de rejeter l'hypothèse que les valeurs de deux KPI ou KQI sont identiques alors qu'elles le sont effectivement.

Le test de signification statistique permet de conclure qu'un opérateur fonctionne mieux qu'un autre (dans le cas de l'évaluation comparative concurrentielle) et de déterminer si une nouvelle technologie ou fonction apporte une amélioration notable (dans le cas de l'évaluation comparative interne).

Outre la signification statistique, il faut impérativement utiliser des seuils de différence pertinente spécifiques aux KPI ou aux KQI lorsque les différences ne présentent pas de pertinence ou se situent dans la plage de précision de mesure de chaque KPI ou KQI. Les seuils de différence pertinente (THrelv) spécifiques aux KPI ou aux KQI sont définis dans d'autres Recommandations de l'UIT-T de la série consacrée à l'évaluation comparative.

Le Tableau 2 donne un exemple de la méthode de comparaison des KPI ou des KQI à utiliser. Afin de maintenir le caractère général de la présente Recommandation, la comparaison porte sur les indicateurs KQI1 et KQI2, à titre d'exemple. KQI1 et KQI2 peuvent donc être n'importe quel KQI choisi pour un service donné (on peut également se servir de KPI). On calcule les valeurs d'écart moyen et type pour KQI1 et KQI2. L'utilisation du nombre d'échantillons disponibles et l'application des équations A-1 et A-2 du test des hypothèses montrent que la qualité des réseaux 1 et 2 pour les indicateurs KQI1 et KQI2 est statistiquement identique. Il apparaît en outre que la différence pour chacun des KQI, KQI1 et KQI2, demeure inférieure à THrelv. La qualité de fonctionnement des réseaux 1 et 2 est statistiquement différente dans le cas du troisième indicateur, KQI3. Mais la différence (0,02) demeurant inférieure à THrelv (0,025), il ne faut pas conclure que la qualité de fonctionnement des réseaux 1 et 2 est différente.

**Tableau 2 – Exemple de signification statistique**

Service	KQI	Réseau 1			Réseau 2			Statistique à NC 95% (Z>1,96)	StatDiff	THrelv
		moyen	type	N	moyen	type	N			
Service 1	KQI1	3,27	0,3	287	3,35	0,6	212	1,78	Non	0,09
	KQI2	0,02	0,14	12	0,015	0,12	10	0,09	Non	0,006
	KQI3	0,93	0,26	69	0,91	0,29	71	2,04	Oui	0,025

Ce type d'analyse appliqué à des résultats d'évaluation comparative détaillés peut être élargi à divers services ainsi qu'à un ensemble plus large de KPI ou de KQI par service, comme mentionné au § 8. En outre, comme le décrit le § 9.4, il est possible de classer des KPI ou des KQI donnés sur plusieurs réseaux sur la base des résultats de la signification statistique (statistique Z à NC de 95%, Tableau 2).

Il convient de noter qu'avancer qu'une configuration de réseau ou de service est "meilleure" qu'une autre requiert, outre la signification statistique, la définition d'un seuil de différence pertinente spécifique à un KPI ou un KQI ainsi que des informations sur la précision des mesures, comme défini dans d'autres Recommandations de l'UIT-T de la série consacrée à l'évaluation comparative.

### 9.3.4 Rapport de résultats

L'analyse statistique et les résultats de l'évaluation comparative doivent être consignés dans un rapport avec la description détaillée des scénarios d'essai et des conditions de l'évaluation comparative, sinon l'interprétation des résultats risque d'être erronée et par conséquent inutile.

### 9.4 Notation et classement des KPI ou des KQI

Il est possible de noter et de classer les KPI ou les KQI de bout en bout dans chacune des zones objets de la campagne d'évaluation comparative. En outre, comme mentionné au § 9.3.3, il faut s'appuyer sur la signification statistique pour être en mesure de faire la distinction entre les différences de qualité de fonctionnement fines, de plus en plus fréquentes entre les réseaux existants des opérateurs. Il faut recourir aux seuils de différence pertinente spécifiques des KPI ou des KQI si les différences ne sont pas pertinentes ou se situent dans la plage de précision des mesures décrite dans d'autres Recommandations de l'UIT-T de la série consacrée à l'évaluation comparative.

La méthodologie de notation et de classement statistiques d'un jeu de KPI ou de KQI fait l'objet de l'Annexe B et le Tableau 3 en fournit un exemple.

**Tableau 3 – Exemple de notation et de classement de KPI ou de KQI de bout en bout statistiquement significatifs**

	Réseau 1				Réseau 2				
	KPI/KQI	type	N	StatDiff	KPI/KQI	type	N	StatDiff	THrelv
KPI1/KQI1	0,95	0,22	87	0,05	0,97	0,17	69	0,00	0,018
KPI2/KQI2	0,93	0,26	87	0,00	0,91	0,29	69	0,23	0,019
KPI3/KQI3	3,89	0,50	2 600	0,00	3,56	0,70	2 070	17,15	0,31
KPI4/KQI4	105,00	5,00	435	42,67	70,00	15,00	350	0,00	34
KPI5/KQI5	1 200,00	300,00	87	0,00	1 800,00	275,00	69	12,31	596

Le Tableau 3 fournit les valeurs cumulées pour la même zone pendant la même période dans le but d'une comparaison à l'identique (voir les instructions du Tableau 1). Outre les valeurs de qualité de fonctionnement, les écarts types sont calculés et le nombre d'échantillons est indiqué. La valeur

StatDiff (voir l'Annexe B) est calculée pour chaque KPI ou KQI sous forme de différence statistiquement significative par rapport au KPI ou au KQI le plus performant (surligné en jaune dans le Tableau 3) à un niveau de confiance de 95% (voir l'Annexe B). Plus la valeur StatDiff est faible, plus la valeur est proche du KPI ou du KQI le plus performant, représentée par StatDiff = 0. En outre, il convient de noter que, dans tous les cas, les différences entre les KQI des deux réseaux sont supérieures à THrev1, ce qui signifie que l'analyse de signification statistique prévaut.

Le classement des KPI ou des KQI basé sur cette analyse apparaît dans le Tableau 4.

**Tableau 4 – Exemple de classement statistique de KPI ou de KQI**

<b>KPI/KQI</b>	<b>Réseau 1</b>	<b>Réseau 2</b>
KPI1/KQI1	Rang 2	Rang 1
KPI2/KQI2	Rang 1	Rang 2
KPI3/KQI3	Rang 1	Rang 2
KPI4/KQI4	Rang 2	Rang 1
KPI5/KQI5	Rang 1	Rang 2

Ce classement peut être élargi à un jeu plus important de KPI ou de KQI en vue d'une évaluation comparative détaillée ou interne.

Dans certains cas, la note statistique du réseau par service et pour tous les services pourra être requise. L'Annexe A présente une technique possible d'exécution de ce calcul. À noter cependant que ce type de technique n'est valable que si elle est entièrement décrite et basée sur des suppositions et des conditions étayées techniquement.

## Annexe A

### Signification statistique à appliquer dans l'analyse de l'évaluation comparative des réseaux mobiles

(Cette Annexe fait partie intégrante de la présente Recommandation.)

On entend par analyse comparative la comparaison entre des KPI ou des KQI décrivant la qualité de fonctionnement des réseaux de divers opérateurs. Pour être utile, la comparaison doit s'appuyer sur des tests de signification statistique (tests d'hypothèses) dépendant des types de KPI ou de KQI comparés, des paramètres permanents (par exemple, MOS, paramètres RF) et des ratios (par exemple, ratios d'aboutissement ou d'échec).

Dans le premier cas, l'équation A-1 calcule la différence significative [UIT-T P.1401]:

$$Z = \text{StatDiff} / \sqrt{\text{std1}^2 / N1 + \text{std2}^2 / N2} > Z_{th} \quad (\text{A-1})$$

où StatDiff indique la différence entre les indicateurs comparés, leurs écarts types std1 et std2 et le nombre total d'échantillons (N1 et N2) utilisés pour comparer chaque indicateur. En d'autres termes, si Z est supérieur à Zth (sur la base d'une distribution gaussienne pour plus de 30 échantillons, à un niveau de confiance NC%), StatDiff est une différence statistiquement significative au niveau de confiance NC%.

Dans le second cas, le type de ratio de KPI ou de KQI est décrit par p, le nombre de succès ou d'échecs sur le nombre total d'échantillons. La différence significative est donnée par l'équation A-2 [b-UIT-T P.1401].

$$Z = \text{StatDiff} / \sqrt{p1*(1-p1)/N1 + p2*(1-p2)/N2} > Z_{th} \quad (\text{A-2})$$

où p1 et p2 représentent le nombre de succès ou d'échecs pour chacun des indicateurs comparés.

Le Tableau A.1 met en correspondance les seuils de signification Zth à différents niveaux de confiance.

**Tableau A.1 – Mise en correspondance des seuils de signification à différents niveaux de confiance**

NC%	90	95	96	97	98	99
Zth	1,64	1,96	2,05	2,17	2,33	2,58

Si moins de 30 échantillons sont disponibles, il faut utiliser une distribution t-Student dans laquelle t-Student (n) où n = N-1 est le nombre de degrés de liberté, pour un nombre total N d'échantillons.

À noter que, outre la signification statistique, il faut impérativement utiliser des seuils de différence pertinente spécifiques aux KPI ou aux KQI lorsque les différences ne présentent pas de pertinence ou se situent dans la plage de précision de mesure de chaque KPI ou KQI. Les seuils de différence pertinente (THrelv) spécifiques aux KPI ou aux KQI sont définis dans d'autres Recommandations de l'UIT-T de la série consacrée à l'évaluation comparative.

## Annexe B

### Notation et classement statistiques de la qualité de fonctionnement d'un réseau

(Cette Annexe fait partie intégrante de la présente Recommandation.)

La présente Annexe décrit l'algorithme de notation et de classement de la qualité de fonctionnement de bout en bout des réseaux utilisés dans les calculs du Tableau 3.

- Calculer les KPI de bout en bout (ou les KQI) du service analysé pour chaque réseau ou opérateur:
  - KPI<sub>1</sub>...KPI<sub>i</sub>...KPI<sub>N</sub>, i=1,n peut être soit une moyenne, soit une médiane soit des proportions (ratios).
- Construire une matrice d'évaluation comparative pour j=1, réseaux M (opérateurs) par service décrit par les indicateurs N KPI ou KQI - voir le Tableau B.1.

**Tableau B.1 – Matrice d'évaluation comparative**

	Netwk_1.....	Netwk_j.....	Netwk_M
KPI_1...	KPI_1,1	KPI_1,j	KPI_1,M
KPI_i..	KPI_i,1	KPI_i,j	KPI_i,M
KPI_N	KPI_N,1	KPI_N,j	KPI_N,M

- Calculer la différence de signification statistique pour chaque KPI<sub>i,j</sub> de la matrice.
 

Pour i=1,N

  - Sélectionner la meilleure valeur KPI<sub>i,best</sub>, "best" étant le réseau j présentant la meilleure qualité de fonctionnement.
  - Calculer la différence de signification statistique StatDiff<sub>i,j</sub> par KPI<sub>i,j</sub> par rapport à la meilleure valeur KPI<sub>i,best</sub> fondée sur les équations B-1 et B-2 selon le type d'indicateur.
$$\text{StatDiff}_{i,j} = \max \{0, (KPI_{i,best} - KPI_{i,j}) / \sqrt{(\text{std1}^2/N1 + \text{std2}^2/N2)} - Z_{th}\} \quad (\text{B-1})$$

$$\text{StatDiff}_{i,j} = \max \{0, (KPI_{i,best} - KPI_{i,j}) / \sqrt{p1*(1-p1)/N1 + p2*(1-p2)/N2} - Z_{th}\} \quad (\text{B-2})$$

NOTE – Z<sub>th</sub> est F(0,05, N1, N2), le résultat statistique à 95% de signification avec degrés de liberté N1 et N2.

Fin
- Déterminer le réseau ou l'opérateur de rang 1 ("meilleure qualité de fonctionnement") pour le service testé dans la zone.
 

Le rang 1 est attribué au réseau présentant la distance de qualité statistiquement significative minimale pour le service testé sur tous les réseaux ou pour tous les opérateurs.

Pour j=1,M

$$\text{Rank } 1 = \text{Rank}(j) \text{ if } \text{StatDiffQuality\_min} = \min(j=1,M) \{ \text{SUM}(i=1,N) \{ \text{StatDiff}_{i,j} * w_i \} \}$$

où les valeurs StatDiff<sub>i,j</sub> sont telles que définies dans l'équation B-1 ou B-2 et où w<sub>i</sub> représente les pondérations préétablies par service, si l'on souhaite utiliser des pondérations. Sinon, il est possible de recourir à des pondérations unitaires égales.

Fin

Déterminer le rang de tous les autres réseaux ou opérateurs inclus dans la campagne d'évaluation comparative.

Pour j=1,M

$$\text{If } \text{Dist}(j) = \max(0, \text{StatDiffQuality} / \text{StatDiffQuality\_min} - Z_{th}) = 0$$

Rank (j)= Rang 1

(détermine tous les réseaux présentant le même résultat statistique de rang 1)

Ou bien

Ordre ascendant Dist(j)

Rank (j) = position dans le vecteur Dist(j)

Fin

À noter que le classement doit s'appliquer sur la base de la signification statistique ainsi qu'avec les seuils de différence pertinente  $TH_{relv}$  spécifiques aux KPI ou aux KQI, lorsque les différences ne sont pas pertinentes ou se situent dans la plage de précision de mesure de chaque KPI ou KQI. Les seuils de différence pertinente ( $TH_{relv}$ ) spécifiques aux KPI ou aux KQI sont définis dans d'autres Recommandations de l'UIT-T de la série consacrée à l'évaluation comparative.

## Appendice I

### Technique possible de notation et de classement statistiques d'un réseau

(Cet appendice ne fait pas partie intégrante de la présente Recommandation.)

Il peut s'avérer souhaitable de déterminer la note statistique globale d'un réseau. Dans ce but, la qualité de fonctionnement du réseau par type de service sert souvent de critère sous-jacent. Cette note doit être calculée pour chaque zone et cumulée pour les régions, comme l'illustre la Figure 1.

La note statistique par service peut être définie par tous les indicateurs de bout en bout KPI<sub>i</sub> ou KQI<sub>i</sub> (i=1,N) influant sur la qualité de bout en bout globale du service analysé. Elle peut donc être fournie par la somme pondérée de la valeur StatDiff<sub>i</sub> (voir l'Annexe B) de chaque KPI par rapport à ceux présentant la meilleure qualité de fonctionnement, comme décrit au § 9.4. Si elle est disponible, la valeur StatDiff<sub>i</sub> est corrigée en termes de valeurs de seuil de différence pertinente. Le résultat final, le StatScore, décrit la qualité de fonctionnement de bout en bout des réseaux comparés par rapport au réseau présentant la meilleure qualité de fonctionnement.

$$\text{StatScore} = \sum(w_i * \text{StatDiff}_i)$$

Ici, w<sub>i</sub> est la pondération allouée à chaque indicateur KPI ou KQI contribuant à la qualité de service. Plus le StatScore est faible, plus la qualité de fonctionnement est bonne (ou proche de celle du réseau présentant la meilleure qualité de fonctionnement) et plus le rang correspondant est élevé.

Le Tableau I.1 est une nouvelle version du Tableau 3, dans laquelle des exemples de pondérations ont été ajoutés. Ils ne sont fournis qu'à titre indicatif car la définition des pondérations ne relève pas du champ d'application de la présente Recommandation. Les pondérations peuvent également être unitaires s'il est décidé que tous les KPI ou KQI présentent la même importance dans la note statistique globale de la qualité de fonctionnement du réseau. Il convient cependant de noter que, même en cas de recours à des pondérations unitaires égales, si le nombre de KPI ou de KQI augmente ou diminue, la note statistique du réseau peut changer et livrer des résultats statistiques différents.

Par conséquent, la notation et le classement statistiques d'un réseau dans la présente Recommandation n'est valable qu'avec une description et une justification détaillées du choix des KPI ou des KQI et de leurs pondérations sélectionnées sous-jacentes. Sans cette transparence, la notation et le classement statistiques des réseaux ne sont pas valables.

Dans l'exemple du Tableau I.1, le réseau 1 reçoit le meilleur rang 1 avec une note minimum de 2,15, sur la base des pondérations indiquées. À noter en outre que, dans tous les cas, les différences entre les KPI ou les KQI provenant des deux réseaux sont supérieures au THrevl approprié. Autrement dit, les résultats de l'analyse de signification statistique sont valables.

**Tableau I.1 – Exemple de notation et de classement statistiques**

	Réseau 1				Réseau 2				
	KPI	type	N	StatDiff	KPI	type	N	StatDiff	THrevl
KPI1/KQI1	0,95	0,22	87	0,05	0,97	0,17	69	0,00	0,018
KPI2/KQI2	0,93	0,26	87	0,00	0,91	0,29	69	0,23	0,019
KPI3/KQI3	3,89	0,50	2 600	0,00	3,56	0,70	2 070	17,15	0,31
KPI4/KQI4	105,00	5,00	435	42,67	70,00	15,00	350	0,00	34

**Tableau I.1 – Exemple de notation et de classement statistiques**

	Réseau 1				Réseau 2				
KPI5/KQI5	1 200,00	300,00	87	0,00	1 800,00	275,00	69	12,31	596
StatScore				2,15				5,83	
Rang				1				2	

Une campagne d'évaluation comparative complète peut cibler le calcul d'une note statistique globale pour chaque réseau testé pour tous les services pris en charge ( $j = 1, M$  nombre de services pris en charge). À noter qu'une note globale de ce type doit être convenablement justifiée et que les pondérations des services doivent être explicitement transparentes pour se conformer aux prescriptions de la présente Recommandation.

Cette note peut être déterminée par l'addition de toutes les notes statistiques par service calculées selon l'Annexe B et comme montré dans l'exemple du Tableau I.1. Dans ce cas, le réseau présentant la meilleure qualité de fonctionnement obtient une note statistique globale de zéro. Plus la note est faible, meilleure est la qualité de fonctionnement du réseau correspondant telle que définie par l'instanciation de la pondération et de l'addition des KPI ou des KQI par service. La note résultant de cette opération sert souvent à représenter le classement global de réseaux comparés par rapport à celui présentant la meilleure qualité de fonctionnement.

$$\text{GlobalNetScore} = \Sigma(\text{Wser}_i * \text{StatScore}_j)$$

La définition des pondérations  $\text{Wserv}_j$  pour le GlobalNetScore ne relève pas du champ d'application de la présente Recommandation. En revanche, celle-ci fournit des indications à ce sujet.

Selon les politiques ou le but commercial de l'opérateur, le GlobalNetScore peut être calculé sous la forme d'une somme pondérée, similaire aux pondérations utilisées pour le calcul du StatScore. Les pondérations, ainsi que le nombre ou la nature des services à prendre en compte dans le calcul du GlobalNetScore, peuvent être laissés au choix ou à la décision des opérateurs dans le cas du scénario d'évaluation comparative interne (Figure 1).

Alternativement, les pondérations peuvent être décidées sur la base de l'analyse des statistiques d'utilisation (par exemple, données recueillies auprès de multiples utilisateurs), des types de services utilisés et du pourcentage d'utilisation pour divers types de zones.

Cependant, la validité de la note globale dépend de la description et de la justification détaillées du choix des KPI ou des KQI et de leurs pondérations sélectionnées sous-jacentes par service ainsi que de la pondération de chaque service dans la note globale. Sans cette transparence, la note et le classement globaux d'un réseau ne sont pas valables. En outre, comme indiqué dans le texte principal de la présente Recommandation, les notes StatDiff sous-jacentes doivent être considérées à la lumière de la pertinence des différences entre services pour les divers KPI et KQI et mises à 0 si une différence est susceptible d'être significative, mais cela ne contribue pas à une meilleure qualité de fonctionnement du réseau du point de vue de l'utilisateur.

## Bibliographie

- [b-UIT-T E.802] Recommandation UIT-T E.802 (2007), *Cadre et méthode de détermination et d'application de paramètres de qualité de service.*
- [b-UIT-T P.10] Recommandation UIT-T P.10/G.100 (2017), *Vocabulaire relatif à la qualité de fonctionnement et à la qualité de service.*
- [b-UIT-T P.863] Recommandation UIT-T P.863 (2018), *Évaluation objective de la qualité d'écoute perçue.*
- [b-UIT-T P.1203] Recommandation UIT-T P.1203 (2017), *Évaluation paramétrique, fondée sur le flux binaire, de la qualité des services de streaming audiovisuel adaptatif ou à téléchargement progressif, avec transport fiable.*
- [b-UIT-T P.1401] Recommandation UIT-T P.1401 (2012), *Méthodes, paramètres et procédures pour l'évaluation, la qualification et la comparaison statistiques de modèles de prédiction objective de la qualité.*
- [b-ETSI TR 102 581] ETSI TR 102 581, V1.2.1 (2015), *Speech processing, transmission and quality aspects (STQ); A study on the minimum additional required attenuation on the antenna path of the field test equipment.*  
[https://www.etsi.org/deliver/etsi\\_tr/102500\\_102599/102581/01.02.01\\_60/tr\\_102581v010201p.pdf](https://www.etsi.org/deliver/etsi_tr/102500_102599/102581/01.02.01_60/tr_102581v010201p.pdf)
- [b-Mehta] Mehta, S. (2014). *Statistics topics*. CreateSpace. 160 p.
- [b-Shaum] Spiegel, M.R., Schiller, J.J., Srinivasan, R.A. (2013). *Schaum's outlines: Probability and statistics*, 4ème édition. New York, NY: McGraw-Hill. 424 p.

## SÉRIES DES RECOMMANDATIONS UIT-T

Série A	Organisation du travail de l'UIT-T
Série D	Principes de tarification et de comptabilité et questions de politique générale et d'économie relatives aux télécommunications internationales/TIC
<b>Série E</b>	<b>Exploitation générale du réseau, service téléphonique, exploitation des services et facteurs humains</b>
Série F	Services de télécommunication non téléphoniques
Série G	Systèmes et supports de transmission, systèmes et réseaux numériques
Série H	Systèmes audiovisuels et multimédias
Série I	Réseau numérique à intégration de services
Série J	Réseaux câblés et transmission des signaux radiophoniques, télévisuels et autres signaux multimédias
Série K	Protection contre les perturbations
Série L	Environnement et TIC, changement climatique, déchets d'équipements électriques et électroniques, efficacité énergétique; construction, installation et protection des câbles et autres éléments des installations extérieures
Série M	Gestion des télécommunications y compris le RGT et maintenance des réseaux
Série N	Maintenance: circuits internationaux de transmission radiophonique et télévisuelle
Série O	Spécifications des appareils de mesure
Série P	Qualité de transmission téléphonique, installations téléphoniques et réseaux locaux
Série Q	Commutation et signalisation et mesures et tests associés
Série R	Transmission télégraphique
Série S	Équipements terminaux de télégraphie
Série T	Terminaux des services télématiques
Série U	Commutation télégraphique
Série V	Communications de données sur le réseau téléphonique
Série X	Réseaux de données, communication entre systèmes ouverts et sécurité
Série Y	Infrastructure mondiale de l'information, protocole Internet, réseaux de prochaine génération, Internet des objets et villes intelligentes
Série Z	Langages et aspects généraux logiciels des systèmes de télécommunication